Heriot-Watt University
Research Gateway

A Tractable Forward-Backward CPHD Smoother

Citation for published version:
Nagappa, S, Delande, ED, Clark, DE & Houssineau, J 2017, 'A Tractable Forward-Backward CPHD Smoother', *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 1, pp. 201-217. https://doi.org/10.1109/TAES.2017.2649978

Digital Object Identifier (DOI):
10.1109/TAES.2017.2649978

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Publisher's PDF, also known as Version of record

Published In:
IEEE Transactions on Aerospace and Electronic Systems

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

A Tractable Forward–Backward CPHD Smoother

SHARAD NAGAPPA
EMMANUEL D. DELANDE
DANIEL E. CLARK
JÉRÉMIE HOUSSINEAU
Heriot-Watt University, Edinburgh U.K.

To circumvent the intractability of the usual Cardinalized Probability Hypothesis Density (CPHD) smoother, we present an approximate scheme where the population of targets born until and after the starting time of the smoothing are estimated separately and where smoothing is only applied to the estimate of the former population. The approach is illustrated through the implementation of a tractable approximation of the usual CPHD smoother.

Manuscript received January 5, 2014; revised December 3, 2014, July 16, 2015, and January 30, 2016; released for publication August 25, 2016. Date of publication January 9, 2017; date of current version April 17, 2017.

DOI. No. 10.1109/TAES.2017.2649978

Refereeing of this contribution was handled by Ba-Ngu Vo.

This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) Platform Grant (EP/J015180/1), in part by the EPSRC Grant EP/K014277/1 and Grant EP/H010866/1, and in part by the U.K. MoD University Defence Research Collaboration in Signal Processing. The work of J. Houssineau was supported by the Ph.D. scholarship sponsored by DCNS and a tuition fee scholarship sponsored by Heriot-Watt University.

Authors' addresses: The authors are with Heriot-Watt University, Edinburgh EH14 4AS, U.K., E-mail: (snagappa@gmail.com; E.D.Delande@hw.ac.uk; D.E.Clark@hw.ac.uk; j.houssineau@hw.ac.uk).

0018-9251/16/$26.00 © 2017 CCBY

I. INTRODUCTION

Estimation of dynamical systems from sensor data relies on the concepts of stochastic filtering, prediction, and smoothing. The state of the system, of interest to some operator, is unknown and only partially observed through a sensor. Stochastic modeling allows the operator to describe the uncertainty of their knowledge on the system state, the inaccuracies of the sensor, and how the observations produced by the sensor relate to the system state.

Prediction, filtering, and smoothing are distinct components of the problem of sequential estimation of a dynamic system conditioned on noisy observations. Prediction provides a means of forecasting the state of the system at some point in the future conditioned on measurements collected up to the present time by the operator, while filtering aims at estimating the state of the system at regular time steps based on all the measurements collected up to that instant. Smoothing [1]–[6] refers to the estimation of the system state at some *past* time t conditioned on a set of observations collected up to the *present* time $t' > t$. The production of the smoothed estimate is, per construction, delayed in time; however, because it exploits observations posterior to the time of estimation, the smoothed estimate provides a more refined description of the system state than the corresponding filtered estimate.

Smoothing is generally a computationally expensive operation. In this respect, advances in sequential Monte Carlo (SMC) methods [7]–[9] have allowed for more efficient implementations of SMC smoothers [10]–[13].

Multi-object estimation requires extending these methods in order to deal with false measurements (clutter), missed detections, as well as uncertainty in the number of targets. A solution to the problem of multi-object estimation can be found in a direct generalisation of the Bayes filter to multi-object systems using the finite set statistics (FISST) framework. Due to the complexity of the multi-object Bayes filter, Mahler proposed approximate filtering solutions [14].

The Probability Hypothesis Density (PHD) filter [15] estimates the posterior multitarget density using its first moment (also referred to as intensity function or PHD). In Mahler's approach, the multitarget state is represented by a random finite set (RFS), and the filter prediction and update equations are derived following the rules provided by the FISST framework. This principled approach to the multi-object estimation problem allows the PHD filter to account for target birth, missed detections as well as clutter. Additionally, it spares the necessity of a data association step. Following the derivation of the PHD filter, Mahler proposed the Cardinalized PHD (CPHD) filter [16], which propagates a cardinality distribution on the number of targets in addition to the intensity, thereby, improving the quality of the estimation.

More recently, smoothing in the context of the PHD filter received some attention. The PHD smoother and its implementation have been addressed by various authors [17]–[28]. As an example of practical application, multi-object smoothing has recently been applied to per-

son tracking in a crowded environment [22]. It has been shown that, compared to the PHD filter, the PHD smoother reduces the localisation error of the state estimates but is not always able to improve the cardinality estimate, i.e., the estimation of the number of targets present in the scene. In particular, it has been shown that the PHD smoother does not perform well when the probability of target death is significant, leading to an overestimated cardinality. Additionally, missed detections propagate backwards in the PHD smoother, further degrading the cardinality estimate in some cases.

The general multi-object smoothing approach developed using FISST estimates the whole population of targets with a single RFS, thereby, incorporating the targets born *before* and *after* the past time step k where the smoothing process starts. The incorporation of the latter population, however, makes the approach intractable when applied to the CPHD filter because of the presence of target birth in the smoothing step [21]. In this paper, we propose an alternative to the general smoothing approach in which the population of targets born before and after the starting date of the smoothing process are estimated separately in order to circumvent the intractability of the CPHD smoothing step, and we derive a tractable approximation of the usual CPHD smoother. An SMC implementation of the proposed solution is then given, and its performance is compared with the CPHD filter and the usual PHD smoother. We show that the CPHD smoother addresses the drawbacks of the usual PHD smoother by providing stable estimates in the number of targets.

In the following section, we introduce multi-object filtering and the proposed multi-object smoothing approach, circumventing the intractability of the general approach when applied to the CPHD filter. Section III summarizes the CPHD filter and we derive a tractable approximation of the usual CPHD smoother in Section IV. An SMC implementation of the filter and smoother is presented in Section V, and we analyze the performance of the algorithm in Section VI.

II. MULTI-OBJECT FILTERING AND SMOOTHING

A. Point Processes

In multiple target detection and tracking problems, the objects of interest—the *targets*—have individual states in some target space $\mathcal{X} \subset \mathbb{R}^d$ that describes some characteristics relevant to the operator (e.g., position and velocity coordinates). Just as the target states, the number of targets is unknown and possibly time varying. In the multi-object filtering framework, the knowledge of the operator about the whole target population is represented by a *single* random object called a (simple finite) point process, whose number of elements *and* element states are random. A realisation of a (target) point process Φ is a set $\{x_1, \ldots, x_N\}$ of points in $\mathcal{X}$ depicting a specific multitarget configuration.

A simple finite point process Φ can be described by its *probability distribution* p_Φ on the measurable space $(\mathbb{X}, \mathbf{B}_\mathbb{X})$, where $\mathbb{X} = \bigcup_{n\geq 0} \mathcal{X}^n$ is the point process state space, i.e., the space of all the finite vectors $\varphi = (x_1, \ldots, x_N)$ of points in $\mathcal{X}$, and $\mathbf{B}_\mathbb{X}$ is the Borel σ-algebra on $\mathbb{X}$ [29]. The corresponding *probability density* describes the point process as well, and is defined as a symmetric function that vanishes if at least two of its arguments are equal.[1]

Similarly to usual random variables, it is also convenient to describe a point process through its *moment measures*. For any region $B \in \mathbf{B}_\mathcal{X}$, the first moment measure μ_Φ of a point process Φ is given by

$$\mu_\Phi(B) = \mathbb{E}\left[\sum_{x\in\Phi} 1_B(x)\right] \tag{1a}$$

$$= \sum_{n\geqslant 0} \int \left(\sum_{1\leqslant i\leqslant n} 1_B(x_i)\right) p_\Phi(\mathrm{d}x_{1:n}) \tag{1b}$$

where $x_{1:n} = (x_1, \ldots, x_n)$, and 1_B is the indicator function[2] on B. The scalar $\mu_\Phi(B)$ provides the expected number of targets or *mean target number* inside B [14]. The first moment measure plays a central role in the derivation of multi-object filtering algorithms, for it is a meaningful statistics describing the point process that is far less expensive to propagate than the full probability distribution (more details are given later). The corresponding density is called the first moment density or *intensity* or *probability hypothesis density* of the point process; we shall denote it by the same notation μ_Φ when there is no ambiguity.

B. Multi-object Bayes Filter

Within the Bayesian paradigm the law $p_{k|k}$ of the filtered state—i.e., the probability distribution of the point process $\Phi_{k|k}$ representing the population of targets—can be estimated at some time $k \geq 0$ given the sequence of measurement sets $Z_{0:k} = (Z_{k'})_{k'\leq k}$, where $Z_{k'}$ denotes the (possibly empty) set of measurements collected by the operator at time $0 \leq k' \leq k$, each measurement $z \in Z_{k'}$ belonging to some observation space $\mathcal{Z} \subset \mathbb{R}^m$. The multi-object Bayes filter [14], [32] is given by the combination of a one-step prediction and update[3]

$$p_{k|k-1}(\mathrm{d}\xi|Z_{0:k-1}) = \int t_{k|k-1}(\mathrm{d}\xi|\varphi) p_{k-1|k-1}(\mathrm{d}\varphi|Z_{0:k-1}), \tag{2}$$

$$p_{k|k}(\mathrm{d}\xi|Z_{0:k}) = \frac{g_k(Z_k|\xi) p_{k|k-1}(\mathrm{d}\xi|Z_{0:k-1})}{\int g_k(Z_k|\varphi) p_{k|k-1}(\mathrm{d}\varphi|Z_{0:k-1})} \tag{3}$$

[1] The FISST methodology for target tracking [14], [30], [31] considers the representation of point processes through a *multi-object density* f_Φ defined on *sets*, whose elements are per construction *unordered* and *distinct*, and relies on the notion of the *set integral* for the derivation of the filtering algorithms such as the PHD [15] or CPHD [16] filters. In this context, a simple finite point process is also called an RFS. In this paper, we follow the alternative measure-theoretical formulation exploited in [32] and [33], which is based on more general representations of point processes [29], [34], [35].

[2] For a measurable subset $B \in \mathbf{B}_\mathcal{X}$, the indicator function 1_B is defined as the function on $\mathcal{X}$ such that $1_B(x) = 1$ if $x \in B$, $1_B(x) = 0$ otherwise.

[3] When p and q are measures on some common space, we use the notation $p(\mathrm{d}x) = q(\mathrm{d}x)$ to indicate that, for every bounded measurable function f on the same space, it holds that $\int f(x)p(\mathrm{d}x) = \int f(x)q(\mathrm{d}x)$.

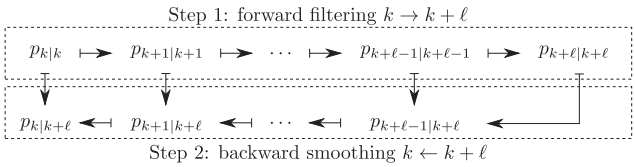

Fig. 1. Forward–backward smoothing for a fixed window $[k, \ldots, k+\ell]$ (general approach). During the forward filtering step, the posterior distributions $p_{k|k}, \ldots, p_{k+\ell|k+\ell}$ are propagated with a multi-object filter (e.g., PHD and Cardinalized PHD (CPHD)). During the backward filtering steps, the smoothed distributions $p_{k+\ell-1|k+\ell}, \ldots, p_{k|k+\ell}$ are produced in reverse order with the corresponding multi-object smoother.

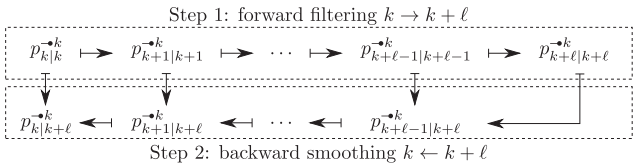

Fig. 2. Forward–backward smoothing for a fixed window $[k, \ldots, k+\ell]$ (proposed approach). During the forward filtering step, the posterior distributions $p_{k|k}^{\rightarrow k}, \ldots, p_{k+\ell|k+\ell}^{\rightarrow k}$ are propagated with an adapted multi-object filter, treating the population of targets born since starting time k as clutter. During the backward filtering steps, the smoothed distributions $p_{k+\ell-1|k+\ell}^{\rightarrow k}, \ldots, p_{k|k+\ell}^{\rightarrow k}$ are produced in reverse order with the corresponding multi-object smoother, in which the modeled birth is removed.

where $t_{k|k-1}$ is the multitarget Markov transition kernel between time steps $k-1$ and k, that incorporates the target motion model, the target birth and death mechanisms, and g_k is the multimeasurement/multitarget likelihood at time step k, that incorporates the sensor observation, the false alarm and missed detection mechanisms.

In the prediction and update equations (2) and (3), the point process $\Phi_{k_1|k_2}$ represents the multitarget configuration in the scene at time $k_1 \geq 0$ given the sequence of measurement sets $(Z_{k'})_{k' \leq k_2}$. For the sake of simplicity, the conditional dependencies on measurement sets in the expression of probability distributions (or densities) involving point processes will be omitted from now on when there is no ambiguity, i.e., $p_{k_1|k_2}(\cdot)$ will denote $p_{k_1|k_2}(\cdot|Z_{0:k_2})$.

C. Multi-object Smoothing: General Approach

It is possible to refine the law $p_{k|k}$ of the filtered state in *past* times exploiting *current* measurements step through a process referred to as smoothing. Given some smoothing lag $\ell > 0$ and some sequence of measurement sets $(Z_{k'})_{k' \leq k+\ell}$, the smoothed distribution $p_{k|k+\ell}$ is obtained using the forward–backward recursion [21], for every time step k' from $k+\ell-1$ down to k as

$$p_{k'|k+\ell}(\mathrm{d}\xi) = \int t_{k'|k'+1}(\mathrm{d}\xi|\varphi) p_{k'+1|k+\ell}(\mathrm{d}\varphi) \qquad (4)$$

where $t_{k'|k'+1}$ is the kernel defined by

$$t_{k'|k'+1}(\mathrm{d}\xi|\varphi) = \frac{t_{k'+1|k'}(\varphi|\xi) p_{k'|k'}(\mathrm{d}\xi)}{\int t_{k'+1|k'}(\varphi|\zeta) p_{k'|k'}(\mathrm{d}\zeta)} \qquad (5)$$

where the multitarget Markov transition kernel introduced in (2) is assumed to have a density, denoted by the same notation $t_{k'+1|k'}$. Note the similarities between the construction of the posterior distribution $p_{k|k}$ in (3) and the kernel $t_{k'|k'+1}$ through (5), for both follow Bayes' rule. This will be exploited for the derivation of the CPHD smoother in Section IV.

The general structure of a multi-object forward–backward smoother on the time window $[k, \ldots, k+\ell]$ is illustrated in Fig. 1 and can be split into the following two steps.

1) *Forward filtering* Produce the posterior distributions $p_{k|k}, p_{k+1|k+1}, \ldots, p_{k+\ell|k+\ell}$ using (2) and (3).
2) *Backward smoothing* Produce the smoothed distributions $p_{k+\ell-1|k+\ell}, p_{k+\ell-2|k+\ell}, \ldots, p_{k|k+\ell}$ using (4) and (5).

D. Multi-object Smoothing: Proposed Approach

This paper follows an alternative approach to the previous approach presented in Section II-C. From the previous approach to multi-object smoothing, we may observe that, for a smoothing step on a fixed window $[k, \ldots, k+\ell]$, the posterior distributions $p_{k|k}, p_{k+1|k+1}, \ldots, p_{k+\ell|k+\ell}$ propagated by the forward filtering step describe the process Φ representing the *whole* population of targets, including those born since the starting time k (that is, the multitarget Markov transition kernel in the prediction equation (2) incorporates a birth target process representing the newborn targets in the whole population). Since the presence of this birth target process makes the corresponding CPHD smoother intractable [21], we propose an alternative approach separating the targets born before and after the starting time k, leading to an approximate but tractable CPHD smoother.

For the rest of this paper, we denote by $\Phi^{-\bullet k}$ (respectively (resp.) $\Phi^{k\circ-}$) the process representing the population of targets born *until* time k (resp. *since* time k, k excluded). The general structure of the proposed multi-object forward–backward smoother on the time window $[k, \ldots, k+\ell]$ is illustrated in Fig. 2 and can be split into the following two steps.

1) *Forward filtering* Produce the posterior distributions $p_{k|k}^{\rightarrow k}, p_{k+1|k+1}^{\rightarrow k}, \ldots, p_{k+\ell|k+\ell}^{\rightarrow k}$ of the process $\Phi^{-\bullet k}$, treating the population of targets born since k as clutter.
2) *Backward smoothing* Produce the smoothed distributions $p_{k+\ell-1|k+\ell}^{\rightarrow k}, p_{k+\ell-2|k+\ell}^{\rightarrow k}, \ldots, p_{k|k+\ell}^{\rightarrow k}$ using adapted smoothing equations in which the modeled birth is removed.

In order to proceed with the aforementioned first point, note that *both* the population of targets born until *and* since time k need to be estimated: The former is the output of the forward filtering step, while the latter is used to model clutter. The proposed forward filtering approach is detailed in the next section.

E. Forward Filtering Without Birth: Principle

Following the principle exposed in Section II-D, the forward filtering step aims at providing the posterior distributions of the target population process $\Phi^{\rightarrow\bullet k}$, representing the targets born until time k. In order to provide better estimates the processes $\Gamma_{k+1}, \ldots, \Gamma_{k+\ell}$, representing the modeled newborn targets in the subsequent time steps relevant to the smoothing window, are not discarded but rather treated as clutter, i.e., as undesirable sources of observations. Essentially, the proposed forward filtering step is composed of two filters running in parallel.

1) The primary filter propagates the posterior distribution $p^{\rightarrow\bullet k}$ of the process $\Phi^{\rightarrow\bullet k}$ (i.e., representing the population of targets born *until* starting time k).
2) The secondary filter propagates the posterior distribution $p^{k\circ-}$ of the process $\Phi^{k\circ-}$ (i.e., representing the population of targets born *since* time k, excluded).

Following the notations in [36], for some process Φ on the target state space $\mathcal{X}$, we denote by $\Theta_k(\Phi)$ the conditional process on the observation space $\mathcal{Z}$ defined by

$$\Theta_k(\Phi) = \bigcup_{x\in\Phi} \theta_k(x) \tag{6}$$

where $\theta_k(x)$ is the Bernoulli process corresponding to the single-measurement/single-target observation model conditional on target state x [see (21)]. Let us fix some time step k' relevant to the smoothing window, i.e., $k \leq k' < k+\ell$. Since the primary filter focusses on the estimation of the targets born *until* time step k, both the targets born *since* time step k and the current false alarms produced by the sensor are undesirable sources of observations. Thus, the update step of the primary filter is fed with the outer population process representing these undesirable sources, i.e.

$$\mathcal{K}^{\rightarrow\bullet k}_{k'+1} = \mathcal{K}_{k'+1} \cup \Theta_{k'+1}(\Phi^{k\circ-}_{k'+1|k'}) \tag{7}$$

where $\Phi^{k\circ-}_{k'+1|k'}$ is the predicted process propagated by the secondary filter, and $\mathcal{K}_{k'+1}$ is the modeled clutter process, accounting for the false alarms produced by the sensor at time step $k'+1$. Conversely, the update step of the secondary filter is fed with the outer population process

$$\mathcal{K}^{k\circ-}_{k'+1} = \mathcal{K}_{k'+1} \cup \Theta_{k'+1}(\Phi^{\rightarrow\bullet k}_{k'+1|k'}) \tag{8}$$

where $\Phi^{\rightarrow\bullet k}_{k'+1|k'}$ is the predicted process propagated by the primary filter. The concept of the proposed forward filtering scheme is illustrated for a generic multi-object filter in Fig. 3.

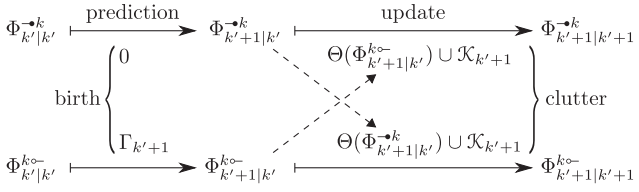

Fig. 3. Forward filtering for a fixed window $[k, \ldots, k+\ell]$, illustrated at some time k', $k \leq k' < k+\ell$ (generic filter). Arrows in plain style denote filtering steps, dotted arrows denote information transfer between filters. The primary filter (above) propagates the law $p^{\rightarrow\bullet k}$ of the population $\Phi^{\rightarrow\bullet k}$, representing the targets born before the beginning of the smoothing window (hence, no birth in the prediction step). The secondary filter (below) propagates the law $p^{k\circ-}$ of the population $\Phi^{k\circ-}$, representing the population of targets born since the beginning of the smoothing window (hence, the modeled target birth $\Gamma_{k'+1}$ in the prediction step). Note from Fig. 2 that only the law $p^{\rightarrow\bullet k}$ is involved in the forward–backward smoothing process.

III. CPHD FILTER

The multi-object Bayes filtering equations (2) and (3)—let alone the smoothing ones (4), (5)—are intractable in the general case, and additional assumptions are necessary in order to derive a practical solution [14]. Both derivations of the two most popular multi-object filters within the FISST framework are based on a central assumption considering the predicted target process $\Phi_{k+1|k}$ as either a *Poisson* or a *independent and identically distributed (i.i.d.)* point process: The former assumption leads to the construction of the PHD filter [15] and the latter to the CPHD filter [16].

A. General Notations

Throughout of this paper we shall use the following notations. The binomial coefficients C_i^j and P_i^j are defined by

$$C_i^j = \begin{cases} \dfrac{j!}{i!(j-i)!}, & 0 \leq i \leq j \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

$$P_i^j = \begin{cases} \dfrac{j!}{(j-i)!}, & 0 \leq i \leq j \\ 0, & \text{otherwise.} \end{cases} \tag{10}$$

We shall also exploit the notations

$$\langle q, \rho \rangle = \sum_{n\geq 0} q(n)\rho(n) \tag{11}$$

$$\rho * \rho' : n \to \sum_{m=0}^{n} \rho(m)\rho'(n-m) \tag{12}$$

$$\mu(f) = \int f(x)\mu(\mathrm{d}x) \tag{13}$$

$$\mu(B) = \int_B \mu(\mathrm{d}x) \tag{14}$$

where ρ and ρ' denote some cardinality distributions, μ some measure, q and f some suitable functions, and B some suitable region. In addition, e_d will denote the elementary symmetric function of order d

$$e_d(\xi) = \sum_{\varphi\subseteq\xi, |\varphi|=d} \left(\prod_{\zeta\in\varphi} \zeta \right) \tag{15}$$

where ξ denotes some (possibly empty) set. Finally, δ will denote, depending on the context, the Kronecker or the Dirac delta function.

B. PHD Filtering and Limitations

A Poisson point process is fully characterized by its first moment measure (1), and the aim of the PHD filter is to propagate the first moment measure (or density) μ_Φ of the target point process Φ within the Bayesian paradigm. Mahler derived the PHD filtering equations in [15]; since then, several implementations of the PHD filter have been proposed [14], [31], [36]–[38]. More recently, adaptations of the smoothing equations (4) and (5) have been investigated and formulations of the corresponding PHD filter with smoothing—or PHD smoother—have been presented [17], [24], [26]–[28].

While inexpensive, the PHD filter has some drawbacks due to the restrictive nature of the Poisson assumption. Because the first and second moment measures of a Poisson process are equal, the variance in the estimated number of targets is significant [16], [39], [40]. In addition, the estimation is generally unstable in the presence of missed detections and high levels of clutter [16].

C. CPHD Filter: Prediction and Update

An i.i.d. point process is fully characterized by its first moment measure (1) *and* its cardinality distribution ρ_Φ, which describes the number of elements in the point process. Accordingly, the CPHD filter jointly propagates the reduced distribution (ρ_Φ, μ_Φ) of the target point process Φ within the Bayesian paradigm. Mahler derived the CPHD filtering equations in [14], but no attempt has been done to produce a corresponding CPHD smoother. In this section, $k > 0$ designs some time step relevant to the scenario.

1) Prediction Equation Denoting by ψ the "empty" state that describes a target that has left the surveillance scene, the single-target transition kernel $t_{k|k-1}$ from the target space $\mathcal{X}$ to the augmented target state space $\{\psi\} \cup \mathcal{X}$ is defined as follows:

$$\begin{cases} t_{k|k-1}(\mathrm{d}x'|x) = p_{\mathrm{s},k}(x)\hat{t}_{k|k-1}(\mathrm{d}x'|x), & x' \in \mathcal{X} \\ t_{k|k-1}(\psi|x) = 1 - p_{\mathrm{s},k}(x) \end{cases} \tag{16}$$

where $\hat{t}_{k|k-1}$ is the single-target Markov transition kernel, from $\mathcal{X}$ to $\mathcal{X}$, describing the evolution of the target states since time $k-1$, and $p_{\mathrm{s},k}$ is the probability of survival of targets since time $k-1$.

THEOREM 1 (PREDICTION [16], [41]) Assuming that the newborn targets can be represented by a point process Γ_k with cardinality distribution $\rho_{\mathrm{b},k}$ and first moment measure γ_k, the predicted cardinality distribution $\rho_{k|k-1}$ and first moment measure $\mu_{k|k-1}$ are found to be

$$\rho_{k|k-1} = \rho_{\mathrm{s},k|k-1} * \rho_{\mathrm{b},k} \tag{17}$$

$$\mu_{k|k-1}(\mathrm{d}x) = \mu_{k-1|k-1}(t_{k|k-1}(\mathrm{d}x|\cdot)) + \gamma_k(\mathrm{d}x) \tag{18}$$

where $\rho_{\mathrm{s},k|k-1}$ is the cardinality distribution of the surviving targets

$$\rho_{\mathrm{s},k|k-1}(i) = r_k^{\mathrm{p}}(\mathcal{X})^i \sum_{j \geq i} C_i^j r_k^{\mathrm{p}}(\psi)^j \rho_{k-1|k-1}(j) \tag{19}$$

with $r_k^{\mathrm{p}}(\psi)$ and $r_k^{\mathrm{p}}(\mathcal{X})$ defined as the ratios

$$\begin{cases} r_k^{\mathrm{p}}(\psi) = \dfrac{\mu_{k-1|k-1}(t_{k|k-1}(\psi|\cdot))}{\mu_{k-1|k-1}(\mathcal{X})} \\ r_k^{\mathrm{p}}(\mathcal{X}) = \dfrac{\mu_{k-1|k-1}(t_{k|k-1}(\mathcal{X}|\cdot))}{\mu_{k-1|k-1}(t_{k|k-1}(\psi|\cdot))}. \end{cases} \tag{20}$$

2) Update Equation Denoting by ϕ the "empty" observation that describes a missed detection, the likelihood g_k in the augmented observation space $\{\phi\} \cup \mathcal{Z}$ is defined as follows:

$$\begin{cases} g_k(z|x) = p_{\mathrm{d},k}(x)\hat{g}_k(z|x), & z \in \mathcal{Z} \\ g_k(\phi|x) = 1 - p_{\mathrm{d},k}(x) \end{cases} \tag{21}$$

where $\hat{g}_k$ is the single-measurement/single-target likelihood in $\mathcal{Z}$ and $p_{\mathrm{d},k}$ is the probability of detection at time k.

THEOREM 2 (UPDATE [16], [32]) Assuming that *1)* the false alarms can be represented by a point process $\mathcal{K}_k$ with cardinality distribution $\rho_{\mathrm{c},k}$, intensity $\mu_{\mathrm{c},k}$, and spatial distribution κ_k, and *2)* the predicted target process $\Phi_{k|k-1}$ is an i.i.d. point process, the updated cardinality distribution $\rho_{k|k}$ and first moment measure $\mu_{k|k}$ are found to be

$$\rho_{k|k}(n) = \frac{\Upsilon_k^0[\mu_{k|k-1}, Z_k](n)\rho_{k|k-1}(n)}{\langle \Upsilon_k^0[\mu_{k|k-1}, Z_k], \rho_{k|k-1} \rangle} \tag{22}$$

$$\mu_{k|k}(\mathrm{d}x) = \left[g_k(\phi|x)\ell_k^{\mathrm{u}}(\phi) + \sum_{z \in Z_k} g_k(z|x)\ell_k^{\mathrm{u}}(z) \right] \mu_{k|k-1}(\mathrm{d}x) \tag{23}$$

where we follow the notation introduced by Vo *et al.* in [40]

$$\Upsilon_k^j[\mu_{k|k-1}, Z](n) = \frac{1}{\mu_{k|k-1}(g_k(\phi|\cdot))^j} \times \sum_{d=0}^{\min(|Z|,n)} \frac{n!(|Z|-d)!}{(n-(d+j))!} \rho_{\mathrm{c},k}(|Z|-d) r_k^{\mathrm{u}}(\phi)^n e_d(Z) \tag{24}$$

with e_d applied to some set $\{\frac{r_k^{\mathrm{u}}(z)}{\kappa_k(z)} | z \in Z\}$ referred as $e_d(Z)$ for notational convenience, and with $r_k^{\mathrm{u}}(\phi)$, $r_k^{\mathrm{u}}(z)$ defined as the ratios

$$\begin{cases} r_k^{\mathrm{u}}(\phi) = \dfrac{\mu_{k|k-1}(g_k(\phi|\cdot))}{\mu_{k|k-1}(\mathcal{X})} \\ r_k^{\mathrm{u}}(z) = \dfrac{\mu_{k|k-1}(g_k(z|\cdot))}{\mu_{k|k-1}(g_k(\phi|\cdot))} \end{cases} \tag{25}$$

and where the corrector terms $\ell_k^{\mathrm{u}}(\phi)$ and $\ell_k^{\mathrm{u}}(z)$ are given by

$$\begin{cases} \ell_k^{\mathrm{u}}(\phi) = \dfrac{\langle \Upsilon_k^1[\mu_{k|k-1}, Z_k], \rho_{k|k-1} \rangle}{\langle \Upsilon_k^0[\mu_{k|k-1}, Z_k], \rho_{k|k-1} \rangle} \\ \ell_k^{\mathrm{u}}(z) = \kappa_k(z)^{-1} \dfrac{\langle \Upsilon_k^1[\mu_{k|k-1}, Z_k \setminus z], \rho_{k|k-1} \rangle}{\langle \Upsilon_k^0[\mu_{k|k-1}, Z_k], \rho_{k|k-1} \rangle}. \end{cases} \tag{26}$$

3) Update Equation Without Clutter In scenarios without clutter, the CPHD update (22), (23) takes a simpler form. Since we require this form of update for the analogy with the smoother in the following section, we present this case here.

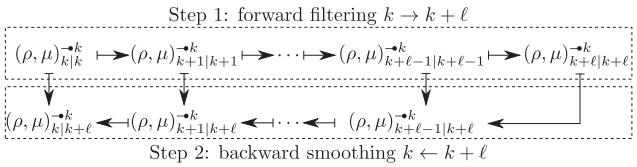

Fig. 4. Forward–backward smoothing for a fixed window $[k,\ldots,k+\ell]$ (CPHD filter). During the forward filtering step, the posterior distributions $(\rho,\mu)^{\rightarrow\bullet k}_{k|k},\ldots,(\rho,\mu)^{\rightarrow\bullet k}_{k+\ell|k+\ell}$ are propagated with an adapted CPHD filter, treating the population of targets born since starting time k as clutter. During the backward filtering steps, the smoothed distributions $(\rho,\mu)^{\rightarrow\bullet k}_{k+\ell-1|k+\ell},\ldots,(\rho,\mu)^{\rightarrow\bullet k}_{k|k+\ell}$ are produced in reverse order with the corresponding CPHD smoother, in which the modeled birth is discarded.

COROLLARY 1 (SIMPLIFIED UPDATE [42]) Under the same assumptions as Theorem 2, and the additional assumption that there is no clutter at time k, the updated cardinality distribution $\rho_{k|k}$ and first moment measure $\mu_{k|k}$ are found to be

$$\rho_{k|k}(n)=\frac{P^n_{|Z_k|}r^{\mathrm{u}}_k(\phi)^n\rho_{k|k-1}(n)}{\sum_{q\geq|Z_k|}P^q_{|Z_k|}r^{\mathrm{u}}_k(\phi)^q\rho_{k|k-1}(q)} \tag{27}$$

$$\mu_{k|k}(\mathrm{d}x)=\left[g_k(\phi|x)\tilde{\ell}^{\mathrm{u}}_k(\phi)+\sum_{z\in Z_k}g_k(z|x)\tilde{\ell}^{\mathrm{u}}_k(z)\right]\mu_{k|k-1}(\mathrm{d}x) \tag{28}$$

where the reduced corrector terms $\tilde{\ell}^{\mathrm{u}}_k(\phi)$ and $\tilde{\ell}^{\mathrm{u}}_k(z)$ are given by

$$\begin{cases}\tilde{\ell}^{\mathrm{u}}_k(\phi)=\dfrac{\sum_{n\geq|Z_k|+1}P^n_{|Z_k|+1}r^{\mathrm{u}}_k(\phi)^n\rho_{k|k-1}(n)}{\mu_{k|k-1}(g_k(\phi|\cdot))\sum_{q\geq|Z_k|}P^q_{|Z_k|}r^{\mathrm{u}}_k(\phi)^q\rho_{k|k-1}(q)}\\ \tilde{\ell}^{\mathrm{u}}_k(z)=\dfrac{1}{\mu_{k|k-1}(g_k(z|\cdot))}.\end{cases} \tag{29}$$

IV. FORWARD–BACKWARD CPHD SMOOTHER

In a similar manner as the PHD and CPHD filters provide a tractable approximation of the multitarget Bayes filter by focusing on the first moment measure of the target point process $\mu_{k|k}$ rather than its full probability distribution $p_{k|k}$, a suitable approximation of the smoothing (4) can be produced through the expression of the first moment $\mu_{k|k+\ell}$ of the smoothed posterior $p_{k|k+\ell}$. Vo *et al.* [24], [28] and Mahler *et al.* [27] developed a PHD smoother using the PHD filter to approximate the forward filter and a backward recursion to approximate the smoothed posterior intensity. In this section, an analogous construction is proposed for the CPHD filter; we use the *simplified* update equation proposed in Section III-C in order to develop a tractable approximation of the usual CPHD smoother.

A. Principle

Following the proposed smoothing approach exposed in Section II-D, the smoothing step in the context of the CPHD filter shall be split into the following two steps (see Fig. 4).

1) *Forward filtering* Produce the posterior distributions $(\rho,\mu)^{\rightarrow\bullet k}_{k|k},(\rho,\mu)^{\rightarrow\bullet k}_{k+1|k+1},\ldots,(\rho,\mu)^{\rightarrow\bullet k}_{k+\ell|k+\ell}$ of the process $\Phi^{\rightarrow\bullet k}$, treating the population of targets born since k as clutter.
2) *Backward smoothing* Produce the smoothed distributions $(\rho,\mu)^{\rightarrow\bullet k}_{k+\ell-1|k+\ell},(\rho,\mu)^{\rightarrow\bullet k}_{k+\ell-2|k+\ell},\ldots,(\rho,\mu)^{\rightarrow\bullet k}_{k|k+\ell}$ using adapted smoothing equations in which the modeled birth is discarded.

We first provide the backward smoothing equations for the CPHD filter in Section IV-B, assuming no target birth between the starting time k and the ending time $k+\ell$. In Section IV-C, we adapt the generic forward filtering approach proposed in Section II-E to the CPHD filter, on some fixed smoothing window $[k,\ldots,k+\ell]$. We then generalize the concept on a sliding window in Section IV-D.

B. Backward Smoothing

In this section, the backward smoothing must be produced on some time window $[k,\ldots,k+\ell]$; we assume that the posterior distribution $(\rho,\mu)^{\rightarrow\bullet k}_{k'|k'}$ and the smoothed distribution $(\rho,\mu)^{\rightarrow\bullet k}_{k'+1|k+\ell}$ are available for some time $k\leq k'<k+\ell$ and we wish to produce the smoothed distribution $(\rho,\mu)^{\rightarrow\bullet k}_{k'|k+\ell}$ (see Fig. 4).

We need to evaluate the smoothed first moment $\mu^{\rightarrow\bullet k}_{k'|k+\ell}$ through (4), assuming that the updated target process $\Phi^{\rightarrow\bullet k}_{k'|k'}$ is i.i.d. Note that the construction of the posterior distribution $p_{k|k}$ in (3) and of the kernel $t_{k|k+1}$ in (5) follow a similar structure through Bayes' rule. Where Corollary 1 allowed the construction of the first moment measure $\mu_{k|k}$ of the posterior distribution $p_{k|k}$ assuming no *clutter*, we can produce an analogous construction of the first moment measure $\tau_{k'|k'+1}$ of the kernel $t_{k'|k'+1}$ assuming no *target birth*:

LEMMA 1 (FORWARD RECURSION) Assuming that 1) the updated target process $\Phi^{\rightarrow\bullet k}_{k'|k'}$ is an i.i.d. point process, and 2) there is no target birth at time k', the cardinality $\tau_{k'|k'+1}$ and first moment measure $\nu_{k'|k'+1}$ of the kernel $t_{k'|k'+1}$ are found to be

$$\tau_{k'|k'+1}(n|\varphi)=\frac{P^n_{|\varphi|}r^{\mathrm{s}}_{k'}(\psi)^n\rho^{\rightarrow\bullet k}_{k'|k'}(n)}{\sum_{q\geq|\varphi|}P^q_{|\varphi|}r^{\mathrm{s}}_{k'}(\psi)^q\rho^{\rightarrow\bullet k}_{k'|k'}(q)} \tag{30}$$

$$\nu_{k'|k'+1}(\mathrm{d}x|\varphi)=\left[t_{k'+1|k'}(\psi|x)\tilde{\ell}^{\mathrm{s}}_{k'}(\psi)+\sum_{y\in\varphi}t_{k'+1|k'}(y|x)\tilde{\ell}^{\mathrm{s}}_{k'}(y)\right]\mu^{\rightarrow\bullet k}_{k'|k'}(\mathrm{d}x) \tag{31}$$

where $r^{\mathrm{s}}_{k'}(\psi)$ is defined as the ratio

$$r^{\mathrm{s}}_{k'}(\psi)=\frac{\mu^{\rightarrow\bullet k}_{k'|k'}(t_{k'+1|k'}(\psi|\cdot))}{\mu^{\rightarrow\bullet k}_{k'|k'}(\mathcal{X})} \tag{32}$$

and where the reduced corrector terms $\tilde{\ell}^{\mathrm{s}}_{k'}(\psi)$ and $\tilde{\ell}^{\mathrm{s}}_{k'}(y)$ are given by

$$\begin{cases} \tilde{\ell}^{\mathrm{s}}_{k'}(\psi) = \dfrac{\sum_{n\geq|\varphi|+1} P^n_{|\varphi|+1} r^{\mathrm{s}}_{k'}(\psi)^n \rho^{-\bullet k}_{k'|k'}(n)}{\mu^{-\bullet k}_{k'|k'}(t_{k'+1|k'}(\psi|\cdot)) \sum_{q\geq|\varphi|} P^q_{|\varphi|} r^{\mathrm{s}}_{k'}(\psi)^q \rho^{-\bullet k}_{k'|k'}(q)} \\ \tilde{\ell}^{\mathrm{s}}_{k'}(y) = \dfrac{1}{\mu^{-\bullet k}_{k'|k'}(t_{k'+1|k'}(y|\cdot))}. \end{cases} \tag{33}$$

The proof is given in Appendix A. The main result of this paper then follows:

THEOREM 3 (CPHD SMOOTHER (TIME k, LAG $\ell > 0$)) Under the same assumptions as Lemma 1, the smoothed cardinality distribution $\rho^{-\bullet k}_{k'|k+\ell}$ and first moment measure $\mu^{-\bullet k}_{k'|k+\ell}$, at some time $k \leq k' < k+\ell$, are found to be

$$\rho^{-\bullet k}_{k'|k+\ell}(n) = \sum_{m\geq 0} \frac{P^n_m r^{\mathrm{s}}_{k'}(\psi)^n \rho^{-\bullet k}_{k'|k'}(n)}{\sum_{q\geq m} P^q_m r^{\mathrm{s}}_{k'}(\psi)^q \rho^{-\bullet k}_{k'|k'}(q)} \rho^{-\bullet k}_{k'+1|k+\ell}(m) \tag{34}$$

$$\mu^{-\bullet k}_{k'|k+\ell}(\mathrm{d}x) = \left[c^{\mathrm{md}}_{k'}(x) + c^{\mathrm{d}}_{k'}(x)\right] \mu^{-\bullet k}_{k'|k'}(\mathrm{d}x) \tag{35}$$

where the missed detection $c^{\mathrm{md}}_{k'}$ and detection $c^{\mathrm{d}}_{k'}$ corrector terms are given by

$$\begin{cases} c^{\mathrm{md}}_{k'}(x) = \dfrac{t_{k'+1|k'}(\psi|x)}{\mu^{-\bullet k}_{k'|k'}(t_{k'+1|k'}(\psi|\cdot))} \\ \qquad \times \sum_{m\geq 0} \dfrac{\sum_{n\geq m+1} P^n_{m+1} r^{\mathrm{s}}_{k'}(\psi)^n \rho^{-\bullet k}_{k'|k'}(n)}{\sum_{q\geq m} P^q_m r^{\mathrm{s}}_{k'}(\psi)^q \rho^{-\bullet k}_{k'|k'}(q)} \rho^{-\bullet k}_{k'+1|k+\ell}(m) \\ c^{\mathrm{d}}_{k'}(x) = \int \dfrac{t_{k'+1|k'}(y|x)}{\mu^{-\bullet k}_{k'|k'}(t_{k'+1|k'}(y|\cdot))} \mu^{-\bullet k}_{k'+1|k+\ell}(\mathrm{d}y). \end{cases} \tag{36}$$

The proof is given in Appendix B.

C. Smoothing on a Fixed Window $[k, \ldots, k+\ell]$

Here, we adapt the generic forward filtering approach presented in Section II-E, i.e., we exploit CPHD filters [16] for the primary and the secondary filters to propagate the posterior distributions $(\rho, \mu)^{-\bullet k}$ and $(\rho, \mu)^{k\circ-}$, respectively (see Fig. 5). As required in the construction CPHD filter [16], the following approximations are necessary, for every time step $k \leq k' < k+\ell$.

(A1) The outer population process[4] $\mathcal{K}^{-\bullet k}_{k'+1}$ is i.i.d.

(A2) The predicted target process $\Phi^{k\circ-}_{k'+1|k'}$ is i.i.d.

(A3) The outer population process $\mathcal{K}^{k\circ-}_{k'+1}$ is i.i.d.

(A4) The predicted target process $\Phi^{-\bullet k}_{k'+1|k'}$ is i.i.d.

PROPOSITION 1 (OUTER POPULATION PROCESSES) Under Approximations IV-C and IV-C, the outer population process $\mathcal{K}^{-\bullet k}_{k'+1}$ is characterized by its distribution $(\rho, \mu)^{-\bullet k}_{\mathrm{c},k'+1}$ given by

$$\begin{cases} \rho^{-\bullet k}_{\mathrm{c},k'+1} = \rho_{\mathrm{c},k'+1} * \rho^{k\circ-}_{\Theta,k'+1|k'} \\ \mu^{-\bullet k}_{\mathrm{c},k'+1} = \mu_{\mathrm{c},k'+1} + \mu^{k\circ-}_{\Theta,k'+1|k'} \end{cases} \tag{37}$$

where $(\rho, \mu)^{k\circ-}_{\Theta,k'+1|k'}$ is the distribution of the process $\Theta_{k'+1}(\Phi^{k\circ-}_{k'+1|k'})$, that is

$$\begin{cases} \rho^{k\circ-}_{\Theta,k'+1|k'}(n) = \left(\dfrac{\mu^{k\circ-}_{k'+1|k'}(p_{\mathrm{d},k'+1}(\cdot))}{\mu^{k\circ-}_{k'+1|k'}(\mathcal{X})}\right)^n \\ \qquad \times \sum_{m\geq n} \rho^{k\circ-}_{k'+1|k'}(m) C^m_n \left(\dfrac{\mu^{k\circ-}_{k'+1|k'}(g_{k'+1}(\phi|\cdot))}{\mu^{k\circ-}_{k'+1|k'}(\mathcal{X})}\right)^{m-n} \\ \mu^{k\circ-}_{\Theta,k'+1|k'}(z) = \mu^{k\circ-}_{k'+1|k'}(g_{k'+1}(z|\cdot)). \end{cases} \tag{38}$$

Likewise, under Approximations IV-C and IV-C, the outer population process $\mathcal{K}^{k\circ-}_{k'+1}$ is characterized by its distribution $(\rho, \mu)^{k\circ-}_{\mathrm{c},k'+1}$ given by (37) and (38) in which "$-\bullet$" subscripts are substituted to "$\circ-$," and vice versa.

The proof is given in Appendix C.

Once all the posterior distributions $(\rho, \mu)^{-\bullet k}_{k+1|k+1}, \ldots, (\rho, \mu)^{-\bullet k}_{k+\ell|k+\ell}$ of the primary CPHD filter are collected through the forward filtering process illustrated in Fig. 5, the smoothed distributions $(\rho, \mu)^{-\bullet k}_{k+\ell-1|k+\ell}, \ldots, (\rho, \mu)^{-\bullet k}_{k|k+\ell}$ are produced as illustrated in Fig. 4, using Theorem 3.

D. Smoothing on a Sliding Window $[k, \ldots, k+\ell]$

Suppose now that we wish to produce the smoothed distributions $(\rho, \mu)^{-\bullet k}_{k|k+\ell}$, for some smoothing lag $\ell > 0$ and for *every* time $k > 0$. Observe in Fig. 5 that the only input for the forward filtering step starting at time k is the filtered distribution $(\rho, \mu)^{-\bullet k}_{k|k}$ of the process $\Phi^{-\bullet k}$ representing all the targets born until time k. Since there is no backward recursion in the usual CPHD filter [16], its output $(\rho, \mu)_{k|k}$ at time k ignores any information acquired in posterior times and *does* describe the process $\Phi^{-\bullet k}$; in other words, the distribution $(\rho, \mu)_{k|k}$ *is* the distribution $(\rho, \mu)^{-\bullet k}_{k|k}$. The

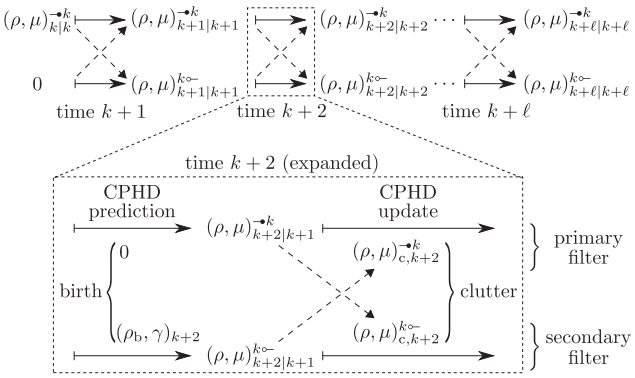

Fig. 5. Forward filtering for a fixed window $[k, \ldots, k+\ell]$ (CPHD filter). Arrows in plain style denote filtering steps, dotted arrows denote information transfer between filters. The primary filter (above) propagates the distribution $(\rho, \mu)^{-\bullet k}$ of the population $\Phi^{-\bullet k}$, representing the targets born before the beginning of the smoothing window (hence, no birth in the prediction steps). The secondary filter (below) propagates the distribution $(\rho, \mu)^{k\circ-}$ of the population $\Phi^{k\circ-}$, representing the population of targets born since the beginning of the smoothing window (hence, the modeled target birth $(\rho_{\mathrm{b}}, \gamma)$ in the prediction steps).

[4]Recall from (7) that the outer population process $\mathcal{K}^{-\bullet k}_{k'+1}$ represents the population of undesirable observations for the primary filter, treated as clutter in the data update step.

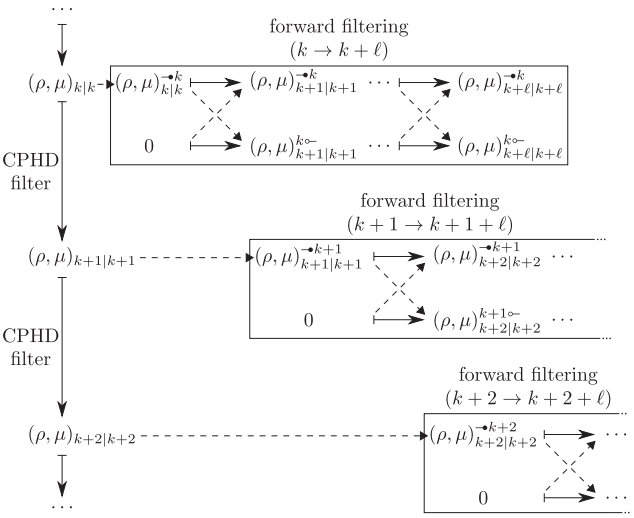

Fig. 6. Forward filtering for a sliding window of fixed smoothing lag $\ell > 0$ (CPHD filter). Arrows in plain style denote filtering steps, dotted arrows denote information transfer between filters. A main CPHD filter (vertical) propagates the posterior distribution of the target process Φ_k. At every time step k, the filtered posterior $(\rho, \mu)_{k|k}$ feeds a new smoothing step over the window $[k, \ldots, k+\ell]$, initialized through an additional pair of primary and secondary CPHD filters.

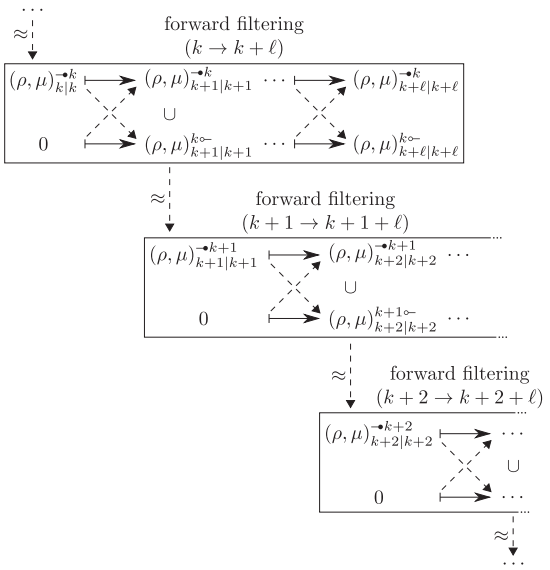

Fig. 7. Simplified forward filtering for a sliding window of fixed smoothing lag $\ell > 0$ (CPHD filter). Arrows in plain style denote filtering steps, dotted arrows denote information transfer between filters. At every time step $k+1$, the outputs $(\rho, \mu)^{\rightarrow\bullet k}_{k+1|k+1}$, $(\rho, \mu)^{k\circ-}_{k+1|k+1}$ of the smoothing step launched at the previous time $k1$ are superposed to form an approximation of the filtered posterior $(\rho, \mu)^{\rightarrow\bullet k+1}_{k+1|k+1}$, the input of the new smoothing step.

CPHD smoother on a sliding window can thus be constructed through a usual CPHD filter, whose output at every time step k feeds the new smoothing step initialized with a new pair of primary and secondary filters (see Fig. 6).

A less expensive alternative, for which the exploitation of the main CPHD filter is spared, can be constructed as follows. Recall that the primary and the secondary filters propagate the law of the processes $\Phi^{\rightarrow\bullet k}_{k'|k'}$ and $\Phi^{k\circ-}_{k'|k'}$, representing the population of targets born *until* and *since* the starting time step k, respectively; essentially, the process $\Phi^{\rightarrow\bullet k}_{k'|k'}$ represents the marginalisation of the process $\Phi_{k'|k'}$ over the targets born since the starting time step k, while the process $\Phi^{k\circ-}_{k'|k'}$ represents the marginalisation of the process $\Phi_{k'|k'}$ over the targets born until the starting time step k. While the superposition of the two processes $\Phi^{\rightarrow\bullet k}_{k'|k'}$ and $\Phi^{k\circ-}_{k'|k'}$ *does not* yield the full target process $\Phi_{k'|k'}$ maintained by the main CPHD filter, in the general case, we may assume the discrepancies to be small enough after a single time step of forward filtering (see Fig. 5) in order to approximate the distribution of the full target process $\Phi_{k+1|k+1}$ as follows.

(A5) For any time k, the distribution $(\rho, \mu)_{k+1|k+1}$ is approximated as

$$\begin{cases} \rho_{k+1|k+1} \approx \rho^{\rightarrow\bullet k}_{k+1|k+1} * \rho^{k\circ-}_{k+1|k+1} \\ \mu_{k+1|k+1} \approx \mu^{\rightarrow\bullet k}_{k+1|k+1} + \mu^{k\circ-}_{k+1|k+1} \end{cases} \tag{39}$$

so that we can exploit the outputs of the first step of the smoothing window that started at time step k to feed the smoothing window starting at the next time step $k+1$, as shown in Fig. 7.

Regardless of the chosen method to initialize the forward filtering processes, i.e., either as illustrated in Fig. 6 or as illustrated in Fig. 7, the backward recursion process is as explained in Section IV-C. That is, once the forward filtering window $(k \to k+\ell)$ is complete, the posterior distributions $(\rho, \mu)^{\rightarrow\bullet k}_{k+1|k+1}, \ldots, (\rho, \mu)^{\rightarrow\bullet k}_{k+\ell|k+\ell}$ of its primary CPHD filter are collected and the smoothed distributions $(\rho, \mu)^{\rightarrow\bullet k}_{k+\ell-1|k+\ell}, \ldots, (\rho, \mu)^{\rightarrow\bullet k}_{k|k+\ell}$ are produced as illustrated in Fig. 4, using Theorem 3.

V. MONTE CARLO IMPLEMENTATION

We summarize in this section the SMC implementation of the CPHD filter as described by Ristic *et al.* [41], [43], as well as an SMC implementation for the CPHD smoother. Equivalent implementations for the PHD filter and smoother have been presented previously [27], [31], [37], [41], [43]. A pseudo code of the SMC-CPHD with smoother follows with Algorithm 1.

A. CPHD Filter

1) *Input* The posterior intensity $\mu_{k-1|k-1}$ is approximated using N_{k-1} particles as

$$\mu_{k-1|k-1} \approx \sum_{n=1}^{N_{k-1}} w^{(n)}_{k-1} \delta_{x^{(n)}_{k-1}}. \tag{40}$$

2) *Prediction* In order to evaluate the predicted cardinality distribution $\rho_{k|k-1}$ with (17), we need to evaluate the ratios (20). Following (40), the numerators and denominators of these ratios can be approximated as follows[5]:

$$\mu_{k-1|k-1}(\mathcal{X}) \approx \sum_{n=1}^{N_{k-1}} w^{(n)}_{k-1} \tag{41}$$

[5]Pay attention to the fact that $\mathcal{X}$ is *not* a variable but denotes the target state space; recall from (14) that $\mu_{k-1|k-1}(\mathcal{X})$ denotes the quantity $\int_{\mathcal{X}} \mu_{k-1|k-1}(\mathrm{d}x)$.

$$\mu_{k-1|k-1}(t_{k|k-1}(\psi|\cdot)) \approx \sum_{n=1}^{N_{k-1}} w_{k-1}^{(n)}\Big[1 - p_{\mathrm{s},k}\big(x_{k-1}^{(n)}\big)\Big] \quad (42)$$

$$\mu_{k-1|k-1}(t_{k|k-1}(\mathcal{X}|\cdot)) \approx \sum_{n=1}^{N_{k-1}} w_{k-1}^{(n)} p_{\mathrm{s},k}\big(x_{k-1}^{(n)}\big). \quad (43)$$

In order to evaluate the predicted intensity $\mu_{k|k-1}$ with (18), we need to construct the target birth intensity γ_k and the predicted intensity $\mu_{\mathrm{s},k|k-1}$ for surviving targets. The target birth intensity is approximated using $N_{\mathrm{b},k}$ uniformly weighted particles drawn following a measurement-driven sampling function b_k [41], i.e.

$$\gamma_k \approx \sum_{n=1}^{N_{\mathrm{b},k}} w_{\mathrm{b},k|k-1}^{(n)} \delta_{x_{\mathrm{b},k}^{(n)}} \quad (44)$$

where

$$\begin{cases} x_{\mathrm{b},k} \sim \mathrm{b}_k(\cdot|Z_k) \\ w_{\mathrm{b},k|k-1}^{(n)} = \dfrac{\nu_{\mathrm{b},k}}{N_{\mathrm{b},k}} \end{cases} \quad (45)$$

with $\nu_{\mathrm{b},k}$ denoting the expected number of newborn targets. Following (40), the predicted intensity $\mu_{\mathrm{s},k|k-1}$ for surviving targets is approximated as

$$\mu_{\mathrm{s},k|k-1} \approx \sum_{n=1}^{N_{k-1}} w_{\mathrm{s},k|k-1}^{(n)} \delta_{x_{\mathrm{s},k|k-1}^{(n)}} \quad (46)$$

where

$$\begin{cases} x_{\mathrm{s},k|k-1}^{(n)} \sim \hat{t}_{k|k-1}\big(\cdot\,|x_{k-1}^{(n)}\big) \\ w_{\mathrm{s},k|k-1}^{(n)} = p_{\mathrm{s},k}\big(x_{k-1}^{(n)}\big) w_{k-1}^{(n)}. \end{cases} \quad (47)$$

Note that the surviving particles are sampled with the transition kernel; while more efficient solutions exist [9], we have chosen to do so for the sake of simplicity.

3) *Update:* Following the adaptive target birth scheme in [41], the weights of the surviving particles $w_{\mathrm{s},k|k-1}^{(n)}$ and those of the newborn particles $x_{\mathrm{b},k|k-1}^{(n)}$ are updated separately.

In order to evaluate the posterior distribution $(\rho, \mu)_{\mathrm{s},k|k}$ of the surviving targets with (22) and (23), we need to evaluate the ratios (25). Following (46), the numerators and denominators of these ratios can be approximated as follows:

$$\mu_{\mathrm{s},k|k-1}(\mathcal{X}) \approx \sum_{n=1}^{N_{k-1}} w_{\mathrm{s},k|k-1}^{(n)} \quad (48)$$

$$\mu_{\mathrm{s},k|k-1}(g_k(\phi|\cdot)) \approx \sum_{n=1}^{N_{k-1}} w_{\mathrm{s},k|k-1}^{(n)}\Big[1 - p_{\mathrm{d},k}\big(x_{\mathrm{s},k|k-1}^{(n)}\big)\Big] \quad (49)$$

$$\mu_{\mathrm{s},k|k-1}(g_k(z|\cdot)) \approx \sum_{n=1}^{N_{k-1}} w_{\mathrm{s},k|k-1}^{(n)} p_{\mathrm{d},k}\big(x_{\mathrm{s},k|k-1}^{(n)}\big)\hat{g}_k\big(z|x_{\mathrm{s},k|k-1}^{(n)}\big). \quad (50)$$

The detection $w_{\mathrm{md},k}^{(n)}$ and missed detection $w_{\mathrm{d},k}^{(n)}$ weight components of each surviving particle $x_{\mathrm{s},k|k-1}^{(n)}$ are kept separate in order to allow for separate resampling process (explained later). That is

$$w_{\mathrm{md},k}^{(n)} = g_k\big(\phi|x_{\mathrm{s},k|k-1}^{(n)}\big)\ell_k^{\mathrm{u}}(\phi) w_{\mathrm{s},k|k-1}^{(n)} \quad (51)$$

$$w_{\mathrm{d},k}^{(n)} = \left[\sum_{z\in Z_k} g_k\big(z|x_{\mathrm{s},k|k-1}^{(n)}\big)\ell_k^{\mathrm{u}}(z)\right] w_{\mathrm{s},k|k-1}^{(n)}. \quad (52)$$

Following the adaptive target birth scheme in [41], the weight of the newborn particles $x_{\mathrm{b},k}^{(n)}$ are updated separately as their probability of detection must be set to one. Therefore, the missed detection term (51) vanishes in the case of newborn particles and the weight update reduces to

$$w_{\mathrm{b},k}^{(n)} = \left[\sum_{z\in Z_k} g_k\big(z|x_{\mathrm{b},k}^{(n)}\big)\ell_k^{\mathrm{u}}(z)\right]_{p_{\mathrm{d},k}=1} w_{\mathrm{b},k|k-1}^{(n)}. \quad (53)$$

4) *Output* Following (23), the updated intensity $\mu_{k|k}$ is thus approximated as

$$\mu_{k|k} \approx \sum_{n=1}^{N_{k-1}} \Big[w_{\mathrm{md},k}^{(n)} + w_{\mathrm{d},k}^{(n)}\Big] \delta_{x_{\mathrm{s},k|k-1}^{(n)}} + \sum_{n=1}^{N_{\mathrm{b},k}} w_{\mathrm{b},k}^{(n)} \delta_{x_{\mathrm{b},k}^{(n)}}. \quad (54)$$

Also, the updated cardinality $\rho_{k|k}$ is evaluated with (22).

5) *Resampling* Resampling needs to be carried out in order to prevent particle degeneracy [9]. Resampling (with replacement) involves replacing particles having low weights with particles, which have a high weight. The resampled particles constitute i.i.d. samples from the posterior with a representation, which is proportional to the weight of the particles.

In the CPHD filter, the intensity associated with the missed detection component is typically low. In a standard resampling scheme, the particles representing the missed detection will be sparsely sampled due to the associated low weights. This increases the probability of track loss due to particle depletion in some regions of the state space [27]. To avoid this, Mahler *et al.* [27] proposed a variation to the resampling scheme for smoothing, whereby, particles corresponding to missed detections are resampled separately. The intensity due to missed detection is approximated by particles with weights $\{w_{\mathrm{md},k}^{(n)}\}_{n=1}^{N_{k-1}}$ while the intensity due to detection is approximated by the same particles with weights $\{w_{\mathrm{d},k}^{(n)}\}_{n=1}^{N_{k-1}}$. By resampling these two sets separately, it becomes possible to retain samples in low intensity regions of the state space, which is critical for good performance of the smoother.

B. CPHD Smoother

Suppose that we wish to produce the smoothed distribution $(\rho, \mu)_{k|k}^{-\bullet k}$ over the fixed window $[k, \ldots, k+\ell]$.

1) *Forward Filtering* The first step consists in producing the posterior distributions $(\rho, \mu)_{k+1|k+1}^{-\bullet k}, \ldots, (\rho, \mu)_{k+\ell|k+\ell}^{-\bullet k}$ of the process $\Phi^{-\bullet k}$ as explained in Section IV-C. As seen in Fig. 5, both the primary and secondary filters are usual CPHD filters that can be implemented following the instructions given in Section V-A.

The only interaction between the primary and secondary CPHD filters occurs following each prediction step, where the predicted distribution $(\rho, \mu)^{k\circ-}_{k'+1|k'}$ maintained by the secondary filter is used to compute the process of indesirable observations of the primary filter (37), and vice versa.

Let us focus on the evaluation of (37); in order to do so, let us evaluate the distribution $(\rho, \mu)^{k\circ-}_{\Theta,k'+1|k'}$ of the process $\Theta_{k'+1}(\Phi^{k\circ-}_{k'+1|k'})$ using (38). Assuming that the predicted intensity $\mu^{k\circ-}_{k'+1|k'}$ maintained by the secondary filter is given by

$$\mu^{k\circ-}_{k'+1|k'} \approx \sum_{n=1}^{N^{k\circ-}_{k'+1|k'}} w^{k\circ-,(n)}_{k'+1|k'} \delta_{x^{k\circ-,(n)}_{k'+1|k'}} \tag{55}$$

then the numerators and denominators in (38) can be approximated as

$$\mu^{k\circ-}_{k'+1|k'}(\mathcal{X}) \approx \sum_{n=1}^{N^{k\circ-}_{k'+1|k'}} w^{k\circ-,(n)}_{k'+1|k'} \tag{56}$$

$$\mu^{k\circ-}_{k'+1|k'}(g_{k'+1}(\phi|\cdot)) \approx \sum_{n=1}^{N^{k\circ-}_{k'+1|k'}} w^{k\circ-,(n)}_{k'+1|k'} \left[1 - p_{\mathrm{d},k'+1}\left(x^{k\circ-,(n)}_{k'+1|k'}\right)\right] \tag{57}$$

$$\mu^{k\circ-}_{k'+1|k'}(p_{\mathrm{d},k'+1}(\cdot)) \approx \sum_{n=1}^{N^{k\circ-}_{k'+1|k'}} w^{k\circ-,(n)}_{k'+1|k'} p_{\mathrm{d},k'+1}\left(x^{k\circ-,(n)}_{k'+1|k'}\right) \tag{58}$$

and the intensity $\mu^{k\circ-}_{\Theta,k'+1|k'}$ can be approximated as

$$\mu^{k\circ-}_{\Theta,k'+1|k'} \approx \sum_{n=1}^{N^{k\circ-}_{k'+1|k'}} w^{k\circ-,(n)}_{\mathrm{o},k'+1|k'} \delta_{z^{k\circ-,(n)}_{k'+1|k'}} \tag{59}$$

where

$$\begin{cases} z^{k\circ-,(n)}_{k'+1|k'} \sim \hat{g}_{k'+1}\left(\cdot\,|x^{k\circ-,(n)}_{k'+1|k'}\right) \\ w^{k\circ-,(n)}_{\mathrm{o},k'+1|k'} = p_{\mathrm{d},k'+1}\left(x^{k\circ-,(n)}_{k'+1|k'}\right) w^{k\circ-,(n)}_{k'+1|k'}. \end{cases} \tag{60}$$

A similar reasoning applies in order to evaluate the distribution $(\rho, \mu)^{-\bullet k}_{\Theta,k'+1|k'}$ of the process $\Theta_{k'+1}(\Phi^{-\bullet k}_{k'+1|k'})$.

2) *Backward Smoothing* We can then produce the smoothed distributions over the window $[k, \ldots, k+\ell]$ using the backward recursion presented in Theorem 3. Suppose that the smoothed intensity $\mu^{-\bullet k}_{k'+1|k+\ell}$ and the posterior intensity $\mu^{-\bullet k}_{k'|k'}$ are approximated by the sets of particles

$$\mu^{-\bullet k}_{k'+1|k+\ell} \approx \sum_{n=0}^{N^{-\bullet k}_{k'+1}} w^{-\bullet k,(n)}_{k'+1|k+\ell} \delta_{x^{-\bullet k,(n)}_{k'+1}} \tag{61}$$

$$\mu^{-\bullet k}_{k'|k'} \approx \sum_{n=0}^{N^{-\bullet k}_{k'}} w^{-\bullet k,(n)}_{k'} \delta_{x^{-\bullet k,(n)}_{k'}} \tag{62}$$

and that we wish to produce the smoothed intensity $\mu^{-\bullet k}_{k'|k+\ell}$ (see Fig. 4). Note that the SMC implementation of the posterior intensity $\mu^{-\bullet k}_{k'|k'}$ and the corresponding smoothed intensity $\mu^{-\bullet k}_{k'|k+\ell}$ use the same set of particles $\{x^{-\bullet k,(n)}_{k'}\}_{n=1}^{N^{-\bullet k}_{k'}}$, for the smoothing process only reweighs the particles without moving them.

Then, the smoothed intensity $\mu^{-\bullet k}_{k'|k+\ell}$ is approximated by the set of particles $\{w^{-\bullet k,(n)}_{k'|k+\ell}, x^{-\bullet k,(n)}_{k'}\}_{n=0}^{N^{-\bullet k}_{k'}}$ where the smoothed weights are given by

$$w^{-\bullet k,(n)}_{k'|k+\ell} = \left[c^{\mathrm{md}}_{k'}\left(x^{-\bullet k,(n)}_{k'}\right) + c^{\mathrm{d}}_{k'}\left(x^{-\bullet k,(n)}_{k'}\right)\right] w^{-\bullet k,(n)}_{k'} \tag{63}$$

where the missed detection c^{md}_k and detection c^{d}_k corrector terms are approximated by

$$\begin{cases} c^{\mathrm{md}}_{k'}(\cdot) \approx \dfrac{t_{k'+1|k'}(\psi|\cdot)}{\sum_{j=1}^{N^{-\bullet k}_{k'}} w^{-\bullet k,(j)}_{k'} t_{k'+1|k'}(\psi|x^{-\bullet k,(j)}_{k'})} \\ \qquad \times \sum_{m\geq 0} \dfrac{\sum_{n\geq m+1} P^n_{m+1} r^{\mathrm{s}}_{k'}(\psi)^n \rho^{-\bullet k}_{k'|k'}(n)}{\sum_{q\geq m} P^q_m r^{\mathrm{s}}_{k'}(\psi)^q \rho^{-\bullet k}_{k'|k'}(q)} \rho^{-\bullet k}_{k'+1|k+\ell}(m) \\ c^{\mathrm{d}}_{k'}(\cdot) \approx \sum_{i=1}^{N^{-\bullet k}_{k'+1}} \dfrac{w^{-\bullet k,(i)}_{k'+1|k+\ell} t_{k'+1|k'}(x^{-\bullet k,(i)}_{k'+1}|\cdot)}{\sum_{j=1}^{N^{-\bullet k}_{k'}} w^{-\bullet k,(j)}_{k'} t_{k'+1|k'}(x^{-\bullet k,(i)}_{k'+1}|x^{-\bullet k,(j)}_{k'})} \end{cases} \tag{64}$$

where the ratio $r^{\mathrm{s}}_{k'}(\psi)$ is approximated by

$$r^{\mathrm{s}}_{k'}(\psi) \approx \frac{\sum_{i=1}^{N^{-\bullet k}_{k'}} w^{-\bullet k,(i)}_{k'} t_{k'+1|k'}\left(\psi\,|x^{-\bullet k,(i)}_{k'}\right)}{\sum_{j=1}^{N^{-\bullet k}_{k'}} w^{-\bullet k,(j)}_{k'}}. \tag{65}$$

Also, the smoothed cardinality $\rho^{-\bullet k}_{k'|k+\ell}$ is evaluated with (34).

VI. RESULTS

In this section, we present results illustrating the performance of the CPHD smoother and comparisons with the PHD smoother. The CPHD filter is used to approximate the forward filter and the backward recursion of the smoothers is applied to this. In particular, Section VI-B1 considers the effect of missed detections on the CPHD smoother, while Section VI-B2 examines the impact of increasing the probability of target death. Section VI-C considers a general multitarget scenario with nonlinear observation model. The optimal subpattern assignment (OSPA) error [44] is used to compare the performance of the CPHD and PHD smoothers.

A. Motion Model and Simulation Parameters

Before discussing the behavior and performance of the CPHD smoother, we present the motion model and other parameters used in the following simulations.

1) *Single Target Motion Model* A constant velocity model is used to describe the single target motion. The single target state is specified as the target in Cartesian coordinates with the associated velocity $x_k = [p_{x,k}, \dot{p_{x,k}}, p_{y,k}, \dot{p_{y,k}}]^T$. The state transition and

Algorithm 1: CPHD Filter and Smoother (time k, lag $\ell > 0$).

Input:

Filtered estimate $\rho_{k-1|k-1}$, $\{(w_{k-1}^{(n)}, x_{k-1}^{(n)})\}_{n=1}^{N_{k-1}}$

Filtering:

Prediction:

Evaluate $\rho_{k|k-1}$ from (17)

Eval. $\{(w_{\mathrm{s},k|k-1}^{(n)}, x_{\mathrm{s},k|k-1}^{(n)})\}_{n=1}^{N_{k-1}}$ from (47)

Update:

Eval. $\rho_{k|k}$ from (22)

Eval. $\{(w_{\mathrm{md},k|k}^{(n)}, w_{\mathrm{d},k|k}^{(n)}, x_{\mathrm{s},k|k-1}^{(n)})\}_{n=1}^{N_{k-1}}$ [see (51), (52)]

Eval. $\{(w_{\mathrm{b},k|k}^{(n)}, x_{\mathrm{b},k})^{(n)}\}_{n=1}^{N_{\mathrm{b},k}}$ from (45), (53)

Resampling:

Sample $\{(w_k^{(n)}, x_k^{(n)})\}_{n=1}^{N_k}$ from $\{(w_{\mathrm{md},k|k}^{(n)}, w_{\mathrm{d},k|k}^{(n)}, x_{\mathrm{s},k|k-1}^{(n)})\}_{n=1}^{N_{k-1}}$ and $\{(w_{\mathrm{b},k|k}^{(n)}, x_{\mathrm{b},k}^{(n)})\}_{n=1}^{N_{\mathrm{b},k}}$

Smoothing:

Forward filtering (from k to $k+\ell$):

Init. primary filter with $\rho_{k|k}$, $\{(w_k^{(n)}, x_k^{(n)})\}_{n=1}^{N_k}$

Init. secondary filter with empty distribution

for $k' = k : k+\ell-1$ **do**

Predition steps (see CPHD filtering aforementioned):

Eval. $\rho_{k'+1|k'}^{-\bullet k}$, $\{(w_{k'+1|k'}^{-\bullet k,(n)}, x_{k'+1}^{-\bullet k,(n)})\}_{n=1}^{N_{k'+1|k'}^{-\bullet k}}$

Eval. $\rho_{k'+1|k'}^{k\circ-}$, $\{(w_{k'+1|k'}^{k\circ-,(n)}, x_{k'+1}^{k\circ-,(n)})\}_{n=1}^{N_{k'+1|k'}^{k\circ-}}$

Primary and secondary clutter:

see (37), (38), (56), and (60)

Update steps (see CPHD filtering aforementioned):

Eval. $\rho_{k'+1|k'+1}^{-\bullet k}$, $\{(w_{k'+1}^{-\bullet k,(n)}, x_{k'+1}^{-\bullet k,(n)})\}_{n=1}^{N_{k'+1}^{-\bullet k}}$

Eval. $\rho_{k'+1|k'+1}^{k\circ-}$, $\{(w_{k'+1}^{k\circ-,(n)}, x_{k'+1}^{k\circ-,(n)})\}_{n=1}^{N_{k'+1}^{k\circ-}}$

end for

Backward recursion (from $k+\ell$ to k):

for $k' = k+\ell-1 : k$ **do**

Eval. smoothed est. $\rho_{k'|k+\ell}$, $\{(w_{k'|k+\ell}^{(n)})\}_{n=1}^{N_{k'}^{-\bullet k}}$ from:

a) Smoothed est. $\rho_{k'+1|k+\ell}$, $\{(w_{k'+1|k+\ell}^{(n)})\}_{n=1}^{N_{k'+1}^{-\bullet k}}$

b) Filtered est. $\rho_{k'}^{-\bullet k}$, $\{(w_{k'}^{-\bullet k,(n)}, x_{k'}^{-\bullet k,(n)})\}_{n=1}^{N_{k'}^{-\bullet k}}$ using (34), (63)

end for

Output:

At time k: filtered est. $\rho_{k|k}$, $\{(w_k^{(n)}, x_k^{(n)})\}_{n=1}^{N_k}$

At time $k+\ell$: smoothed est. $\rho_{k|k+\ell}^{-\bullet k}$, $\{(w_{k|k+\ell}^{-\bullet k,(n)}, x_k^{(n)})\}_{n=1}^{N_k}$

observation models are given by the following equations:

$$x_k = \begin{bmatrix} 1 & \Delta T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta T \\ 0 & 0 & 0 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} \frac{\Delta T^2}{2} & 0 \\ \Delta T & 0 \\ 0 & \frac{\Delta T^2}{2} \\ 0 & \Delta T \end{bmatrix} w_{k-1} \tag{66}$$

where the time between successive observations is $\Delta T = 1$ s and $w_{k-1} \sim \mathcal{N}(0, \sigma_w^2 \mathbb{I}_2)$ is the process noise.

2) *Observation Model* A nonlinear range-bearing observation model is used here. The observation is specified as $z_k = [r_k, \theta_k]^T$ with a the range taking values $r_k \in [0, 2800]$ m and the bearing angle $\theta_k \in [-\pi, \pi]$ rad. Given a state x_k

$$z_k = \begin{bmatrix} \sqrt{p_{x,k}^2 + p_{y,k}^2} \\ \arctan\left(\frac{p_{y,k}}{p_{x,k}}\right) \end{bmatrix} + v_k \tag{67}$$

where $v_k \sim \mathcal{N}(0, \mathrm{diag}([\sigma_r^2, \sigma_\theta^2]^T))$ is assumed to be zero-mean additive white Gaussian noise. The variance on the range and bearing are, respectively, $\sigma_r^2 = 3\,\mathrm{m}^2$ and $\sigma_\theta^2 = 0.035\,\mathrm{rad}^2$. Clutter is modeled as a Poisson process with mean λ_c with a uniform spatial distribution over the observation space.

3) *SMC Implementation* In the SMC implementation used here, 1000 particles are assigned per target. Following the resampling scheme of Mahler *et al.* [27], the particles are equally divided between missed detections and the measurement updated intensity.

The target birth function is modeled as a Gaussian mixture with components centered on the observations [41]; 250 particles are drawn from each component of the mixture with these particles constituting the target birth intensity.

Additionally, we utilize a label-augmented state space to reduce the complexity of the smoother from quadratic to linear in the number of targets [23]. A label τ is assigned to the particle state at the time of target birth such that $\tau^{(i)} = \tau^{(j)}$ if $x^{(i)}$ and $x^{(j)}$ are drawn from the same Gaussian mixture birth component. Then, the augmented-state transition kernel density $\tilde{t}_{k'+1|k'}$ is

$$\begin{aligned} &\tilde{t}_{k'+1|k'}\left((x,\tau)_{k'+1}^{-\bullet k,(i)} \middle| (x,\tau)_{k'}^{-\bullet k,(j)}\right) \\ &\quad = t_{k+1|k}\left(x_{k'+1}^{-\bullet k,(i)} \middle| x_{k'}^{-\bullet k,(j)}\right) \delta_{\tau_{k'+1}^{-\bullet k,(i)}}\left(\tau_{k'}^{-\bullet k,(j)}\right) \end{aligned} \tag{68}$$

so that $\tilde{t}_{k'+1|k'}\left((x,\tau)_{k'+1}^{-\bullet k,(i)} \middle| (x,\tau)_{k'}^{-\bullet k,(j)}\right) = 0$ if $x_{k'+1}^{-\bullet k,(i)}$ and $x_{k'}^{-\bullet k,(j)}$ originate from different birth terms in (64).

4) *Performance Metric* We compare the performance of the algorithms using the optimal sub-pattern assignment (OSPA) metric [44]. The OSPA metric measures the multi-object miss distance between a set of true targets and a set of estimated targets. The distance is composed of a localisation and cardinality error. Here, we use the OSPA metric with parameters cutoff, $c = 100$ m, and power $p = 1$.

B. Two-Target Scenario

Mahler *et al.* showed that the PHD smoother does not necessarily improve cardinality estimates "especially when the birth PHD γ is small compared to the PHD of surviving targets [...] and when the probability of target death $1 - p_S$ is high" [27]. In this section, we show that the CPHD smoother consistently produces more accurate cardinality estimates when compared with the CPHD filter and PHD smoother.

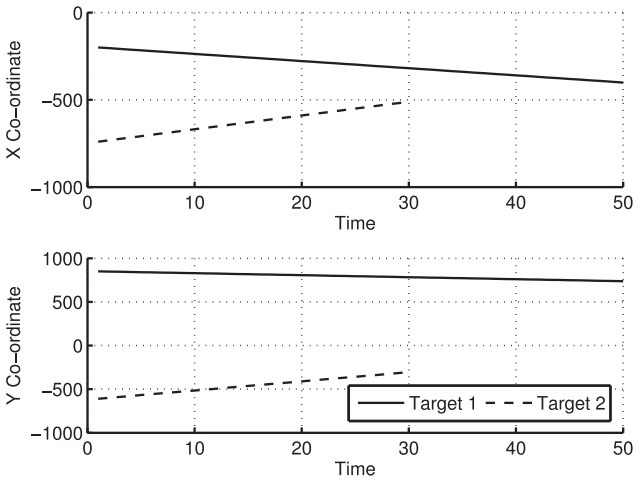

Fig. 8. Solid lines show the true x- and y-coordinates of the two targets in the scene.

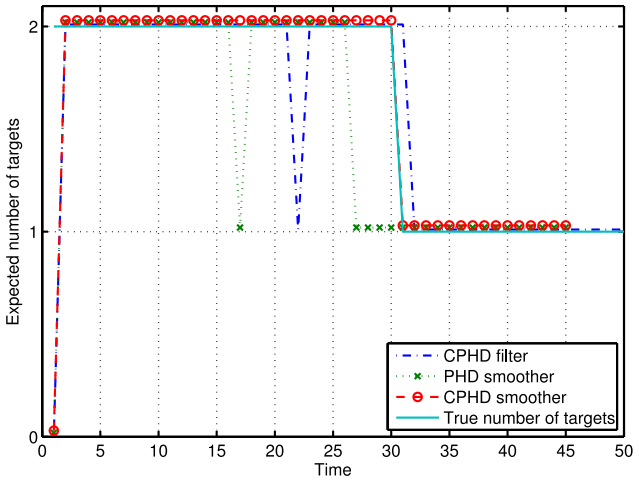

Fig. 9. Mean target number estimated by the CPHD filter, PHD smoother, and CPHD smoother. Target birth occurs only at $k = 1$ and there are no clutter measurements (averaged over 100 Monte Carlo runs).

We consider a tracking scenario involving two targets as shown in Fig. 8. Clutter is not present in this scenario. The CPHD filter is used for the forward filter and fixed lag smoothing with a lag of 5 is performed using the PHD and CPHD smoothers. Target birth in the filter occurs only at the start ($k = 1$) and is centered around the true positions of the targets.

1) *Missed Detections and Target Death* We first examine the effect of missed detection and target death on the smoothers. This simulation uses the scenario illustrated in Fig. 8 with the constraint that the targets are always detected except at two consecutive time steps, $k = \{21, 22\}$, where only one of the targets is detected. The probability of target detection is assumed as $p_{\mathrm{d}} = 0.98$ and the probability of target survival as $p_{\mathrm{s}} = 0.98$. The results from this simulation are averaged over 100 Monte Carlo runs.

Fig. 9 shows the cardinality estimate produced by the filter and smoothers. The filter operates as expected and always suffers a missed detection at $k = \{21, 22\}$. Mahler *et al.* [27] showed that the PHD smoother is unable to

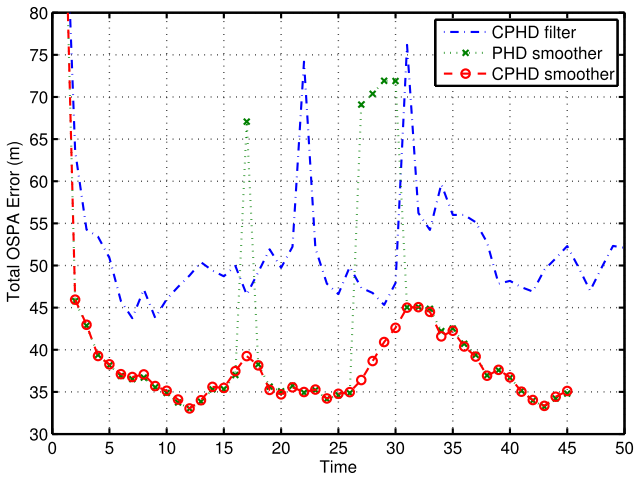

Fig. 10. Total OSPA error ($c = 100$ m, $p = 1$) of the CPHD filter, PHD smoother, and CPHD smoother for a two-target scenario with no clutter (averaged over 100 Monte Carlo runs).

recover tracks lost due to missed detections. Missed detections in the filter propagate backwards through the smoothing recursion resulting in missed detections in the smoother at previous time instants. This missed detection in the filter affects the cardinality estimate of the PHD smoother, which shows a cardinality error at $k = 17$. Mahler *et al.* suggested that this is due to the "strictly first-order nature of the PHD, which propagates cardinality information with a single parameter" [27]. Missed detections in the filter do not propagate backwards in the CPHD smoother due to its more robust formulation.

Fig. 9 also illustrates premature target death, which occurs in the PHD smoother. This is related to the problem of missed detections discussed previously. Target 2 is lost by the CPHD filter at time $k = 32$. This leads to premature target death in the PHD smoother at $k = 27$. The effect of premature death is not seen in the CPHD smoother.

The average total OSPA error of the filter and smoothers is shown in Fig. 10. The figure shows that the PHD and CPHD smoothers exhibit near-identical localisation error. The OSPA error at $k \in \{17, 21, 22, 25 \ldots 31\}$ arises from cardinality error in the PHD smoother. Improvements in the CPHD smoother, thus, manifest as improved cardinality estimates.

2) *Effect of Probability of Survival* We now examine the effect of the probability of target survival p_{s} on the intensity of the smoother. Mahler *et al.* [27] noted that the cardinality estimates of the smoother may not improve when $1 - p_{\mathrm{s}}$ is not negligible. We show that the intensity function of the CPHD smoother is unaffected by the value of p_{s}. For the two-target scenario in Fig. 8, we consider $p_{\mathrm{s}} = \{0.99, 0.98\}$.

Figs. 11 and 12 show the intensity from the filter and smoothers for $p_{\mathrm{s}} = 0.99$ and $p_{\mathrm{s}} = 0.98$, respectively. We see that the PHD smoother overestimates the intensity as $1 - p_{\mathrm{s}}$ increases, and in some cases, this may lead to the reporting of false targets. The CPHD smoother provides stable estimates regardless of the probability of target death and is robust with regards to this parameter.

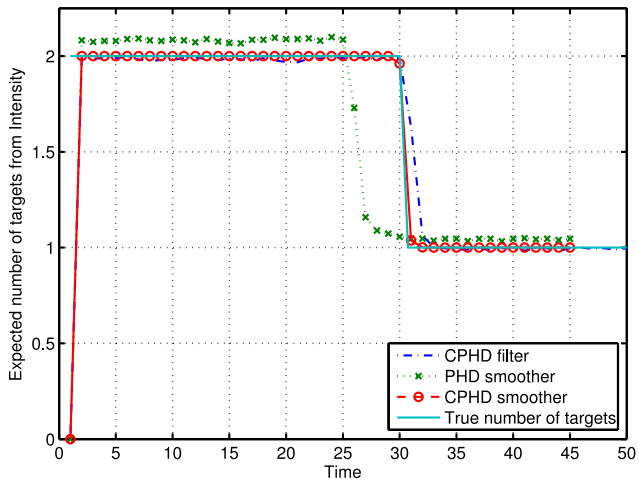

Fig. 11. Mean target number estimated by the CPHD filter, PHD smoother, and CPHD smoother for a two-target scenario when $p_s = 0.99$ (averaged over 100 Monte Carlo runs).

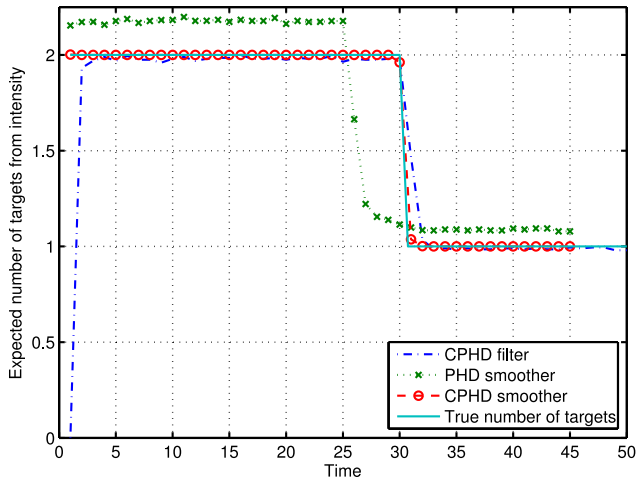

Fig. 12. Mean target number estimated by the CPHD filter, PHD smoother, and CPHD smoother for a two-target scenario when $p_s = 0.98$ (averaged over 100 Monte Carlo runs).

C. Five-Target Scenario

In this section, we consider a more general multitarget filtering and smoothing example where target birth is allowed. The scenario consists of five targets following the trajectories illustrated in Fig. 13. The probability of target detection is $p_d = 0.98$ and the probability of target survival is $p_s = 0.99$. A smoothing lag of 3 is used in these simulations and results are averaged over 500 Monte Carlo runs.

Clutter is modeled as a Poisson process with intensity $\lambda_c = 1.1 \times 10^{-3}$ rad^{-1}m^{-1} over the observation region, resulting in an average of 20 clutter points per scan. Target birth follows a measurement-driven model with an intensity of $\gamma = 0.1$ using 250 particles per measurement.

Figs. 14–16 compare the total (resp., localisation and cardinality) OSPA distance of the filter and smoothers, averaged over 500 Monte Carlo runs. In this general multitarget scenario, the CPHD smoother consistently has a lower OSPA error than the PHD smoother. As discussed

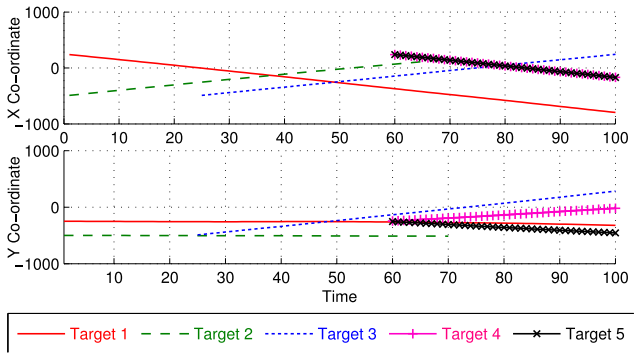

Fig. 13. Multitarget scenario showing the trajectories of five targets in the scene. Clutter intensity is $\lambda_c = 1.1 \times 10^{-3}$ rad^{-1}m^{-1} ($\approx$20 clutter measurements per scan) and measurement-driven target birth intensity is $\gamma = 0.1$.

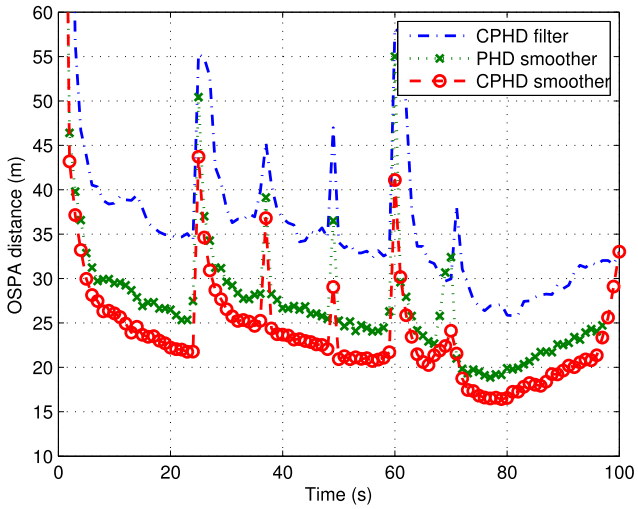

Fig. 14. Total OSPA error ($c = 100$ m, $p = 1$) of the CPHD filter, PHD smoother, and CPHD smoother for a multitarget tracking scenario (averaged over 500 Monte Carlo runs).

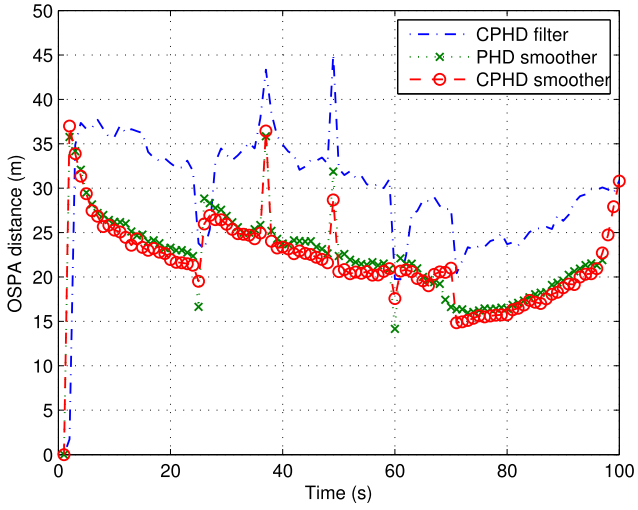

Fig. 15. Localisation OSPA error ($c = 100$ m, $p = 1$) of the CPHD filter, PHD smoother, and CPHD smoother for a multitarget tracking scenario (averaged over 500 Monte Carlo runs).

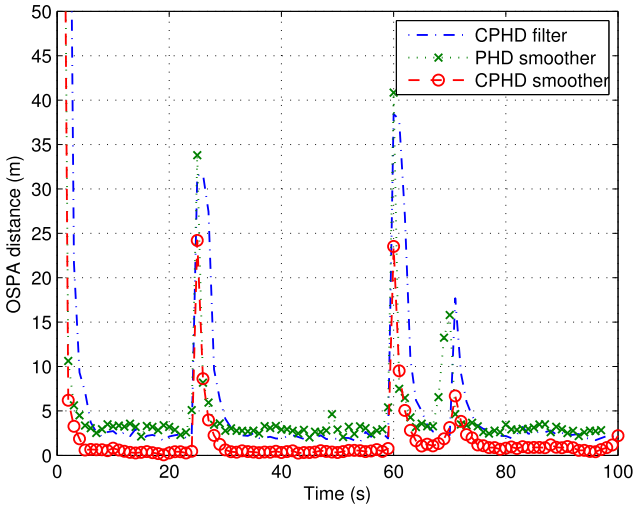

Fig. 16. Cardinality OSPA error ($c = 100$ m, $p = 1$) of the CPHD filter, PHD smoother, and CPHD smoother for a multitarget tracking scenario (averaged over 500 Monte Carlo runs).

previously and shown in Figs. 15 and 16, this is mostly due to improved cardinality estimates of the CPHD smoother, rather than improved localisation.

VII. CONCLUSION

In this paper, a multi-object smoothing approach in which targets born before and after the starting time of the smoothing step are estimated separately is proposed, in order to circumvent the intractability of the general multi-object smoothing approach with FISST when applied to the CPHD filter.

A tractable approximation of the usual CPHD smoother is then derived, and compared with the PHD smoother on a simulation where it is shown to provide better estimates. In particular, the CPHD smoother performance does not deteriorate in the same manner as the PHD smoother when the probability of target death is significant. Additionally, the CPHD smoother is able to improve the cardinality estimate produced by the PHD smoother, and does not exhibit undesirable track deletions as the PHD smoother does.

APPENDIX

A. LEMMA 1

PROOF The proof is straightforward: indeed, considering the structure of the Bayes's rules (3) and (5), the result directly obtained from Corollary 1 in which the parameters of the observation process are replaced by the corresponding parameters in the target evolution process, i.e., $Z_k \to \varphi$, $g_k \to t_{k'+1|k'}$, $\kappa_k \to \gamma_{k'}$, $\rho_{k|k-1} \to \rho^{\rightarrow k}_{k'|k'}$, $\mu_{k|k-1} \to \mu^{\rightarrow k}_{k'|k'}$, $\rho_{k|k} \to \tau_{k|k+1}$, and $\mu_{k|k} \to \nu_{k|k+1}$. ■

B. THEOREM 3

PROOF Let us set some starting time $k \leq 0$, some smoothing lag $\ell > 0$, and some time step $k \leq k' < k + \ell$.

Let us focus first on the cardinality distribution. From the expression of the smoothed distribution $p_{k'|k+\ell}$ in (4), we draw the expression of the smoothed cardinality distribution $\rho_{k'|k+\ell}$, i.e.

$$\rho^{\rightarrow k}_{k'|k+\ell}(n) = \int \tau_{k'|k'+1}(n|\varphi)\, p^{\rightarrow k}_{k'+1|k+\ell}(\mathrm{d}\varphi). \tag{69}$$

We then substitute the result of Lemma 1 in (69) and we get

$$\rho^{\rightarrow k}_{k'|k+\ell}(n) = \int \frac{P^n_{|\varphi|} r^{\mathrm{s}}_{k'}(\phi)^n \rho^{\rightarrow k}_{k'|k'}(n)}{\sum_{q \geq |\varphi|} P^q_{|\varphi|} r^{\mathrm{s}}_{k'}(\phi)^q \rho^{\rightarrow k}_{k'|k'}(q)} p^{\rightarrow k}_{k'+1|k+\ell}(\mathrm{d}\varphi) \tag{70a}$$

$$= \sum_{m \geq 0} \frac{P^n_m r^{\mathrm{s}}_{k'}(\phi)^n \rho^{\rightarrow k}_{k'|k'}(n)}{\sum_{q \geq m} P^q_m r^{\mathrm{s}}_{k'}(\phi)^q \rho^{\rightarrow k}_{k'|k'}(q)} \rho^{\rightarrow k}_{k'+1|k+\ell}(m). \tag{70b}$$

Let us now focus on the first moment measure. From the expression of the smoothed distribution $p_{k'|k+\ell}$ in (4), we draw the expression of the smoothed first moment measure $\mu^{\rightarrow k}_{k'|k+\ell}$, i.e.

$$\mu^{\rightarrow k}_{k'|k+\ell}(\mathrm{d}x) = \int \nu_{k'|k'+1}(\mathrm{d}x|\varphi)\, p^{\rightarrow k}_{k'+1|k+\ell}(\mathrm{d}\varphi). \tag{71}$$

We then substitute the result of Lemma 1 into (71) and we get

$$\mu^{\rightarrow k}_{k'|k+\ell}(\mathrm{d}x) = \left[c^{\mathrm{md}}_{k'}(x) + c^{\mathrm{d}}_{k'}(x)\right] \mu^{\rightarrow k}_{k'|k'}(\mathrm{d}x) \tag{72}$$

with

$$\begin{cases} c^{\mathrm{md}}_{k'}(x) = \int t_{k'+1|k'}(\phi|x) \tilde{\ell}^{\mathrm{s}}_{k'}(\phi)\, p^{\rightarrow k}_{k'+1|k+\ell}(\mathrm{d}\varphi) \\ c^{\mathrm{d}}_{k'}(x) = \int \left(\sum_{y \in \varphi} t_{k'+1|k'}(y|x) \tilde{\ell}^{\mathrm{s}}_{k'}(y) \right) p^{\rightarrow k}_{k'+1|k+\ell}(\mathrm{d}\varphi). \end{cases} \tag{73}$$

Let us focus first on the term $c^{\mathrm{md}}_{k'}$. Substituting the expression of the corrector term $\tilde{\ell}^{\mathrm{s}}_{k'}(\phi)$, given by (33), into (73) yields

$$c^{\mathrm{md}}_{k'}(x) = \frac{t_{k'+1|k'}(\phi|x)}{\mu^{\rightarrow k}_{k'|k'}(t_{k'+1|k'}(\phi|\cdot))} \times \int \frac{\sum_{n \geq |\varphi|+1} P^n_{|\varphi|+1} r^{\mathrm{s}}_{k'}(\phi)^n \rho^{\rightarrow k}_{k'|k'}(n)}{\sum_{q \geq |\varphi|} P^q_{|\varphi|} r^{\mathrm{s}}_{k'}(\phi)^q \rho^{\rightarrow k}_{k'|k'}(q)} p^{\rightarrow k}_{k'+1|k+\ell}(\mathrm{d}\varphi) \tag{74a}$$

$$= \frac{t_{k'+1|k'}(\phi|x)}{\mu^{\rightarrow k}_{k'|k'}(t_{k'+1|k'}(\phi|\cdot))} \times \sum_{m \geq 0} \frac{\sum_{n \geq m+1} P^n_{m+1} r^{\mathrm{s}}_{k'}(\phi)^n \rho^{\rightarrow k}_{k'|k'}(n)}{\sum_{q \geq m} P^q_m r^{\mathrm{s}}_{k'}(\phi)^q \rho^{\rightarrow k}_{k'|k'}(q)} \rho^{\rightarrow k}_{k'+1|k+\ell}(m). \tag{74b}$$

Let us now focus on the term $c^{\mathrm{d}}_{k'}$. Substituting the expression of the corrector term $\tilde{\ell}^{\mathrm{s}}_{k'}(y)$, given by (33) into (73)

yields

$$c^{\mathrm{d}}_{k'}(x) = \int \sum_{y\in\varphi} \frac{t_{k'+1|k'}(y|x)}{\mu^{-\bullet k}_{k'|k'}(t_{k'+1|k'}(y|\cdot))} p^{-\bullet k}_{k'+1|k+\ell}(\mathrm{d}\varphi) \quad (75a)$$

$$= \int \frac{t_{k'+1|k'}(y|x)}{\mu^{-\bullet k}_{k'|k'}(t_{k'+1|k'}(y|\cdot))} \mu^{-\bullet k}_{k'+1|k+\ell}(\mathrm{d}y) \quad (75b)$$

where the last equality was obtained using Campbell's Theorem (see [29], p. 106).

C. PROPOSITION 1

PROOF The time subscripts are omitted in this proof for the sake of clarity.

1) Process $\Theta(\Phi^{k\circ-})$ Under Approximation IV-C, the predicted target process $\Phi^{k\circ-}$ is i.i.d. with distribution $(\rho, \mu)^{k\circ-}$, its probability generating functional (p.g.fl.) [16] $G^{k\circ-}$ is, thus, given by

$$G^{k\circ-}(h) = \sum_{m\geq 0} \rho^{k\circ-}(m) \left(\frac{\mu^{k\circ-}(h)}{\mu^{k\circ-}(\mathcal{X})} \right)^m. \quad (76)$$

Using the notations introduced in (21), the p.g.fl. G_θ of the observation process [16] is given by

$$G_\theta(h|x) = g(\phi|x) + \int h(z) g(z|x) \mathrm{d}z. \quad (77)$$

The p.g.fl. $G^{k\circ-}_\Theta$ of the process $\Theta(\Phi^{k\circ-})$ is then

$$G^{k\circ-}_\Theta(h) = G^{k\circ-}(G_\theta(h|\cdot)) \quad (78a)$$

$$= \sum_{m\geq 0} \rho^{k\circ-}(m) \left(\frac{\mu^{k\circ-}(G_\theta(h|\cdot))}{\mu^{k\circ-}(\mathcal{X})} \right)^m \quad (78b)$$

$$= \sum_{m\geq 0} \rho^{k\circ-}(m) \left(\frac{\mu^{k\circ-}(g(\phi|\cdot))}{\mu^{k\circ-}(\mathcal{X})} + \frac{\int h(z)\mu^{k\circ-}(g(z|\cdot))\mathrm{d}z}{\mu^{k\circ-}(\mathcal{X})} \right)^m \quad (78c)$$

$$= \sum_{m\geq 0} \rho^{k\circ-}(m) \sum_{n=0}^{m} C^m_n \left(\frac{\mu^{k\circ-}(g(\phi|\cdot))}{\mu^{k\circ-}(\mathcal{X})} \right)^{m-n} \times \left(\frac{\int h(z)\mu^{k\circ-}(g(z|\cdot))\mathrm{d}z}{\mu^{k\circ-}(\mathcal{X})} \right)^n \quad (78d)$$

$$= \sum_{n\geq 0} \left(\frac{\int h(z)\mu^{k\circ-}(g(z|\cdot))\mathrm{d}z}{\int \mu^{k\circ-}(g(z|\cdot))\mathrm{d}z} \right)^n \left(\frac{\int \mu^{k\circ-}(g(z|\cdot))\mathrm{d}z}{\mu^{k\circ-}(\mathcal{X})} \right)^n \times \sum_{m\geq n} \rho^{k\circ-}(m) C^m_n \left(\frac{\mu^{k\circ-}(g(\phi|\cdot))}{\mu^{k\circ-}(\mathcal{X})} \right)^{m-n}. \quad (78e)$$

From (78e), we see that the process $\Theta(\Phi^{k\circ-})$ is i.i.d. with cardinality distribution $\mu^{k\circ-}_\Theta$ given by

$$\rho^{k\circ-}_\Theta(n) = \left(\frac{\int \mu^{k\circ-}(g(z|\cdot))\mathrm{d}z}{\mu^{k\circ-}(\mathcal{X})} \right)^n \times \sum_{m\geq n} \rho^{k\circ-}(m) C^m_n \left(\frac{\mu^{k\circ-}(g(\phi|\cdot))}{\mu^{k\circ-}(\mathcal{X})} \right)^{m-n} \quad (79)$$

and intensity function $\mu^{k\circ-}_\Theta$ given by

$$\mu^{k\circ-}_\Theta(z) = \mu^{k\circ-}(g(z|\cdot)). \quad (80)$$

2) Process $\mathcal{K}^{-\bullet k}$ Following its definition (7), the outer population process $\mathcal{K}^{-\bullet k}$ is the superposition of the processes $\mathcal{K}$ and $\Theta(\Phi^{k\circ-})$, which are independent. From [16, eq. (16)], it follows that its cardinality distribution $\rho^{-\bullet k}_{\mathrm{c}}$ is given by

$$\rho^{-\bullet k}_{\mathrm{c}} = \rho_{\mathrm{c}} * \rho^{k\circ-}_\Theta. \quad (81)$$

In addition, its intensity function $\mu^{-\bullet k}_{\mathrm{c}}$ is given by [29, p.152]

$$\mu^{-\bullet k}_{\mathrm{c}} = \mu_{\mathrm{c}} + \mu^{k\circ-}_\Theta. \quad (82)$$

Under Approximation IV-C, the outer population process $\mathcal{K}^{-\bullet k}$ is i.i.d., and is, thus, characterized by its distribution $(\rho, \mu)^{-\bullet k}_{\mathrm{c}}$.

REFERENCES

[1] J. Meditch
A survey of data smoothing for linear and nonlinear dynamic systems
Automatica, vol. 9, no. 2, pp. 151–162, 1973.

[2] B. D. O. Anderson and J. B. Moore
Optimal Filtering. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.

[3] G. Kitagawa
Monte Carlo filter and smoother for nonlinear non-Gaussian state space models
J. Comput. Graphical Statist., vol. 5, no. 1, pp. 1–25, 1996.

[4] Y. Bresler
Two-filter formula for discrete-time non-linear Bayesian smoothing
Int. J. Control, vol. 43, no. 2, pp. 629–641, 1986.

[5] G. Kitagawa
The two-filter formula for smoothing and an implementation of the Gaussian-sum smoother
Ann. Inst. Statistical Math., vol. 46, no. 4, pp. 605–623, 1994.

[6] M. Briers, A. Doucet, and S. Maskell
Smoothing algorithms for state-space models
Ann. Inst. Statistical Math., vol. 62, no. 1, pp. 61–89, 2010.

[7] M. K. Pitt and N. Shephard
Filtering via simulation: Auxiliary particle filters
J. Amer. Statistical Assoc., vol. 94, no. 446, pp. 590–599, 1999.

[8] A. Doucet, N. de Freitas, and N. Gordon
Sequential Monte Carlo Methods in Practice (Statistics for Engineering and Information Science). New York, NY, USA: Springer, 2001.

[9] B. Ristic, S. Arulampalam, and N. Gordon
Beyond the Kalman Filter. Norwood, MA, USA: Artech House, 2004.

[10] A. Doucet, S. J. Godsill, and C. Andrieu
On sequential Monte Carlo sampling methods for Bayesian filtering
Statist. Comput., vol. 18, no. 3, pp. 197–208, 2000.

[11] S. J. Godsill, A. Doucet, and M. West
Monte Carlo smoothing for nonlinear time series
J. Amer. Statistical Assoc., vol. 99, no. 465, pp. 156–168, 2004.

[12] R. Douc, A. Garivier, E. Moulines, and J. Olsson
Sequential Monte Carlo smoothing for general state space hidden Markov models
Ann. Appl. Probability, vol. 21, no. 6, pp. 2109–2145, 2011.

[13] P. Fearnhead, D. Wyncoll, and J. Tawn
A sequential smoothing algorithm with linear computational cost
Biometrika, vol. 97, no. 2, pp. 447–464, 2010.

[14] R. P. S. Mahler
Statistical Multisource-Multitarget Information Fusion. Norwood, MA, USA: Artech House, 2007.

[15] R. P. S. Mahler
Multitarget Bayes filtering via first-order multitarget moments
IEEE Trans. Aerosp. Electron. Syst., vol. 39, no. 4, pp. 1152–1178, Oct. 2003.

[16] R. P. S. Mahler
PHD filters of higher order in target number
IEEE Trans. Aerosp. Electron. Syst., vol. 43, no. 4, pp. 1523–1543, 2007.

[17] N. Nandakumaran, K. Punithakumar, and T. Kirubarajan
Improved multi-target tracking using probability hypothesis density smoothing
In *Proc. SPIE 6699*, Signal and Data Processing of Small Targets 2007, p. 66990M, Sep. 25, 2007, doi:10.1117/12.734656.

[18] D. E. Clark
Joint target-detection and tracking smoothers
In *Proc. SPIE*, vol. 7336, p. 73360G, May 2009.

[19] D. E. Clark, B.-T. Vo, and B.-N. Vo
Forward-backward sequential Monte Carlo smoothing for joint target detection and tracking
In *Proc. 12th Int. Conf. Inf. Fusion*, Jul. 2009, pp. 899–906.

[20] S. I. Hernandez, P. Teal, and M. Frean
Two-filter probability hypothesis density smoothing
In *Proc. First Chilean Workshop Pattern Recog., Theory Appl.*, Nov. 2009, pp. 12–19.

[21] D. E. Clark
First-moment multi-object forward-backward smoothing
In *Proc. 13th Int. Conf. Inf. Fusion*, Jul. 2010, pp. 1–6.

[22] S. I. Hernandez
State Estimation and Smoothing for the Probability Hypothesis Density Filter. Ph.D. thesis, Victoria University of Wellington, Wellington, New Zealand, 2010.

[23] S. Nagappa and D. E. Clark
Fast Monte Carlo PHD smoothing
In *Proc. 14th Int. Conf. Inf. Fusion*, Jul. 2011, pp. 1–7.

[24] B.-N. Vo, B.-T. Vo, and R. P. S. Mahler
A closed form solution to the probability hypothesis density smoother
In *Proc. 13th Int. Conf. Inf. Fusion*, Jul. 2010, pp. 1–8.

[25] B.-T. Vo, D. E. Clark, B.-N. Vo, and B. Ristic
Bernoulli forward-backward smoothing for joint target detection and tracking
IEEE Trans. Signal Process., vol. 59, no. 9, pp. 4473–4477, Sep. 2011.

[26] N. Nadarajah, T. Kirubarajan, T. Lang, M. Mcdonald, and K. Punithakumar
Multitarget tracking using probability hypothesis density smoothing
IEEE Trans. Aerosp. Electron. Syst., vol. 47, no. 4, pp. 2344–2360, Oct. 2011.

[27] R. P. S. Mahler, B.-N. Vo, and B.-T. Vo
Multi-target forward-backward smoothing with the probability hypothesis density
IEEE Trans. Aerosp. Electron. Syst., vol. 48, no. 1, pp. 707–728, Jan. 2012.

[28] B.-N. Vo, B.-T. Vo, and R. P. S. Mahler
Closed-form solutions to forward-backward smoothing
IEEE Trans. Signal Process., vol. 60, no. 1, pp. 2–17, Jan. 2012.

[29] D. Stoyan, W. S. Kendall, and J. Mecke
Stochastic Geometry and Its Applications, 2nd ed. New York, NY, USA: Wiley, 1995.

[30] I. R. Goodman, R. P. S. Mahler, and H.-T. Nguyen
Mathematics of Data Fusion. Norwell, MA, USA: Kluwer, 1997.

[31] B.-N. Vo, S. Singh, and A. Doucet
Sequential Monte Carlo methods for multi-target filtering with random finite sets
IEEE Trans. Aerosp. Electron. Syst., vol. 41, no. 4, pp. 1224–1245, Oct. 2005.

[32] E. D. Delande, M. Uney, J. Houssineau, and D. E. Clark
Regional variance for multiobject filtering
IEEE Trans. Signal Process., vol. 62, no. 13, pp. 3415–3428, Jul. 2014.

[33] M. Andrecki, E. D. Delande, J. Houssineau, and D. E. Clark
Sensor management with regional statistics for the PHD filter
In *Proc. Sensor Signal Process. Defence*, Sep. 2015, pp. 1–5.

[34] D. J. Daley and D. Vere-Jones
An Introduction to the Theory of Point Processes, *vol. 1 ofStatistical Theory and Methods* (Springer Series in Statistics), 2nd ed. New York, NY, USA: Springer, 2003.

[35] D. J. Daley and D. Vere-Jones
An Introduction to the Theory of Point Processes, *vol. 2 ofStatistical Theory and Methods* (Springer Series in Statistics), 2nd ed. New York, NY, USA: Springer, 2008.

[36] B.-N. Vo and W.-K. Ma
The Gaussian mixture probability hypothesis density filter
IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4091–4104, Nov. 2006.

[37] T. Zajic and R. P. S. Mahler
Particle-systems implementation of the PHD multitarget-tracking filter
Proc. SPIE, vol. 5096, pp. 291–299, Aug. 2003.

[38] B.-N. Vo, S. Singh, and A. Doucet
Sequential Monte Carlo implementations of the PHD filter for multi-target tracking
In *Proc. 6h Int. Conf.Inf. Fusion*, Jul. 2003, pp. 792–799.

[39] O. Erdinc, P. Willett, and Y. Bar-Shalom
Probability hypothesis density filter for multitarget multisensor tracking
In *Proc. 8th Int. Conf. Inf. Fusion*, Jul. 2005, pp. 1–8.

[40] B.-T. Vo, B.-N. Vo, and A. Cantoni
Analytic implementations of the cardinalized probability hypothesis density filter
IEEE Trans. Signal Process., vol. 55, no. 7, pp. 3553–3567, Jul. 2007.

[41] B. Ristic, D. E. Clark, B.-N. Vo, and B.-T. Vo
Adaptive target birth intensity for PHD and CPHD filters
IEEE Trans. Aerosp. Electron. Syst., vol. 48, no. 2, pp. 1656–1668, Apr. 2012.

[42] R. P. S. Mahler and V. Maroulas
Tracking spawning objects
IET Radar, Sonar Navigat., vol. 7, no. 3, pp. 321–331, 2013.

[43] B. Ristic, D. E. Clark, and B.-N. Vo
Improved SMC implementation of the PHD filter
In *Proc. 13th Int. Conf. Inf. Fusion*, Jul. 2010, pp. 1–8.

[44] D. Schuhmacher, B.-N. Vo, and B.-T. Vo
A consistent metric for performance evaluation of multi-object filters
IEEE Trans. Signal Process., vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

Sharad Nagappa received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Edinburgh, Edinburgh, U.K., in 2005 and 2010, respectively.

He worked as a Researcher with Heriot-Watt University, Edinburgh, U.K., and the University of Girona, Girona, Spain, until 2013, focusing on multisensor multitarget tracking and its application to simultaneous localisation and mapping for autonomous underwater vehicles. He is currently a Researcher with GE Global Research developing diagnostic solutions for GE Transportation and Oil & Gas businesses, Bangalore, India.

Emmanuel D. Delande received the Eng. degree from the *grande école* Ecole Centrale de Lille, Lille, France, and the M.Sc. degree in automatic control and signal processing from the University of Science & Technology, Lille, both in 2008, and the Ph.D. degree in signal processing, from the Ecole Centrale de Lille, in 2012.

He is currently with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, U.K., where he received a postdoctoral position in the EPSRC-DSTL grant Signal Processing 4 the Networked Battlespace in 2013. His research interests include the design and the implementation of multi-object filtering solutions for multiple target tracking and sensor management problems.

Daniel E. Clark received the Ph.D. degree in signal processing from Heriot-Watt University, Edinburgh, U.K., in 2006.

He is currently an Associate Professor with the School of Engineering and Physical Sciences, Heriot-Watt University. His research interests include the development of the theory and applications of multi-object estimation algorithms for sensor fusion problems. He has led a range of projects spanning theoretical algorithm development to practical deployment.

Jérémie Houssineau received the Eng. degree in mathematical and mechanical modeling from MATMECA, Bordeaux, France, and the M.Sc. degree in mathematical modeling and statistics from the University of Bordeaux, Bordeaux, both in 2009, and the Ph.D. degree in statistical signal processing in 2015 from Heriot-Watt University, Edinburgh, U.K.

He is currently a Research Fellow in the Department of Statistics and Applied Probability, University of Singapore, Singapore. From 2009 to 2011, he was a Research Engineer with DCNS, Toulon, France, involved in the design and assessment of statistical multisensor multi-object estimation algorithms.